

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2002, 14:32:03 ; Search time 67 Seconds
(without alignments)
443.506 Million cell updates/sec

Title: US-09-817-199a-2

Perfect score: 1150

Sequence: 1 MTGPGAVATRDGEAPERSP.....FQIRDYVESQKRSCCSFM 223

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1150	100.0	232	22	AA041696 Human polypeptide
2	1144	99.5	226	22	AA017119 Novel signal trans
3	1133	98.5	222	22	AA017541 Novel signal trans
4	977	85.0	191	22	AAG67156 Amino acid sequenc
5	977	85.0	191	22	AA039910 Human polypeptide
6	977	85.0	191	23	AB076426 Human Ras protein
7	969	84.3	191	23	AA052190 Human degranulatio
8	946	82.3	191	23	AA052189 Mouse degranulatio
9	940	81.7	191	23	AA052188 Degranulation regu
10	926	80.5	191	20	AA030132 An exocytotic prot

11	770	67.0	158	22	AB017244 Novel human diagno
12	697	60.6	666	22	AB071578 Drosophila melanog
13	666	57.9	190	19	AA080747 Human RAB protein,
14	535	46.5	139	22	AA017123 Novel signal trans
15	514.5	44.7	221	23	AB041333 Human ovarian anti
16	513.5	44.7	213	22	AB011916 Human RAB homolog
17	509	44.3	218	21	AA019220 Arabidopsis thalia
18	508.5	44.2	204	22	AB070670 Drosophila melanog
19	507	44.1	253	21	AA039944 Arabidopsis thalia
20	507	44.1	254	21	AA080687 Arabidopsis thalia
21	507	44.1	258	21	AA030498 Arabidopsis thalia
22	505.5	44.0	215	21	AA035215 Zea mays protein f
23	502	43.7	216	21	AA080806 Arabidopsis thalia
24	501	43.6	203	21	AA030499 Arabidopsis thalia
25	501	43.6	207	22	AAG67154 Amino acid sequenc
26	501	43.6	207	22	AA092628 Human protein sequ
27	500	43.5	216	21	AA080688 Arabidopsis thalia
28	500	43.5	216	21	AA039945 Arabidopsis thalia
29	498	43.3	218	21	AA056993 Human prostate can
30	498	43.3	218	22	AB072666 Novel human diagno
31	497.5	43.3	202	21	AA010858 Arabidopsis thalia
32	496.5	43.2	202	21	AA007763 Arabidopsis thalia
33	494	43.0	200	21	AA019165 Amino acid sequenc
34	494	43.0	200	21	AA099979 Human Rab10 protei
35	494	43.0	200	22	AA095340 Human protein sequ
36	488	42.4	224	21	AA047826 Arabidopsis thalia
37	488	42.4	234	21	AA047825 Arabidopsis thalia
38	487.5	42.4	201	21	AA099982 Canine Rab10 prote
39	487.5	42.4	207	22	AB071647 Drosophila melanog
40	487	42.3	209	22	AB023365 Novel human diagno
41	483	42.0	218	21	AA030710 Arabidopsis thalia
42	483	42.0	221	21	AA030709 Arabidopsis thalia
43	483	42.0	246	21	AA058196 Lung cancer associ
44	481.5	41.9	199	21	AA099980 Human Rab10 protei
45	481.5	41.9	199	21	AA099981 Human Rab10 protei

ALIGNMENTS

RESULT 1	
AA041696	
ID	AA041696 standard; Protein; 232 AA.
XX	
AC	AA041696;
XX	
DT	22-OCT-2001 (first entry)
XX	
DE	Human polypeptide SEQ ID NO 6627.
XX	
KW	Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW	peripheral nervous system; neuropathy; central nervous system; CNS;
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW	leukaemia.
XX	
OS	Homo sapiens.
XX	
PN	WO200153312-A1.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000WO-US34263.
XX	
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0598042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.
PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Dmanac RT;
XX WPI; 2001-442253/47.
DR N-PSDB; AAI60852.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Example 2; SEQ ID NO 6627; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAAM2213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 232 AA;
Query Match 100.0%; Score 1150; DB 22; Length 232;
Best Local Similarity 100.0%; Pred. No. 9.6e-120;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTGTPGAVATRDGEAPERSPPSPSYDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTFI 60
DB 10 MTGTPGAVATRDGEAPERSPPSPSYDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTFI 69
QY 61 ATVGIDFRNKVTVGVVRVKLQIWDTAGQERFSVTHAYYRDAQALLLYDITNKSSFDN 120
DB 70 ATVGIDFRNKVTVGVVRVKLQIWDTAGQERFSVTHAYYRDAQALLLYDITNKSSFDN 129
QY 121 IRAWLTIHEYAQRDVVIMLLGNKADMSSEVRSRSEDGETFLAREYGVPPLETSAKTGMNV 180
DB 130 IRAWLTIHEYAQRDVVIMLLGNKADMSSEVRSRSEDGETFLAREYGVPPLETSAKTGMNV 189
QY 181 ELAFLAIKELKYRAGHQADEPSFQIRDYVESQKRSSCCSFM 223
DB 190 ELAFLAIKELKYRAGHQADEPSFQIRDYVESQKRSSCCSFM 232
RESULT 2
AAU17119
ID AAU17119 standard; Protein; 226 AA.
XX
XX AAU17119;
XX
XX 07-NOV-2001 (first entry)
DT
XX Novel signal transduction pathway protein, Seq ID 684.
DE
XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;

KW reproductive system; gastrointestinal; liver disorder; AIDS;
KW acquired immune deficiency syndrome.
XX Homo sapiens.
XX WO200154733-A1.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US01312.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0203515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0218880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.
PR 26-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.

PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465460/50.
DR N-PSDB; AAS27036.
XX Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders -
XX
PS Claim 1; SEQ ID No 684; 880pp; English.
XX The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction
CC pathway protein, amino acid sequences of the invention.
XX

Query Match 99.5%; Score 1144; DB 22; Length 226;

Best Local Similarity 99.6%; Pred. No. 4.3e-119;
Matches 222; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTGTPCAVATRDGEAPERSPSPSYDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTFI 60

Db 4 MTGTPCAVATRDGEAPERSPSPSYDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTFI 63

QY 61 ATVGIDFRNKKVTVVVGVRVKLQIWDTAGQERFSVTHAYYRDAQALLLYDITNKSSFDN 120

Db 64 ATVGIDF\$NKKVTVVVGVRVKLQIWDTAGQERFSVTHAYYRDAQALLLYDITNKSSFDN 123

QY 121 IRAWLTEIHEYAQRDVIIMLGKADMSSEVRIRSEDETGLAREYGVPLETSAKTGMNV 180

Db 124 IRAWLTEIHEYAQRDVIIMLGKADMSSEVRIRSEDETGLAREYGVPLETSAKTGMNV 183

QY 181 ELAFLAI\$KELKYRAGHQADEPSFQIRDYVESQKRSSCCSFM 223

Db 184 ELAFLAI\$KELKYRAGHQADEPSFQIRDYVESQKRSSCCSFM 226

RESULT 3

AAU17541

ID AAU17541 standard; Protein; 222 AA.

XX AC AAU17541;

XX 07-NOV-2001 (first entry)

XX Novel signal transduction pathway protein, Seq ID 1106.

KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;

PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 XX WPI: 2001-465460/50.
 DR N-PSDB; AAS27458.
 XX
 PT Novel polypeptides useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cancers, immune
 PT disorders and neuronal disorders
 XX
 PS Claim 1; SEQ ID No 1106; 880pp; English.
 XX
 CC The invention relates to novel isolated polypeptides (I), and
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
 CC diagnosing, preventing and treating diseases including immune system
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
 CC transplant rejections and graft versus host disease, infectious diseases
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative
 CC disorders, primary haematopoietic disorders, hyperproliferative
 CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
 CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
 CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
 CC Addison's disease), reproductive system disorders, gastrointestinal
 CC disorder (inflammatory disorders), liver disorders (cirrhosis),
 CC as stimulators of B-cell responsiveness to pathogens, activators of
 CC T-cells, to induce higher affinity antibodies, and as a means to induce
 CC tumour proliferation in pathologies e.g. acquired immune deficiency
 CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction
 CC pathway protein, amino acid sequences of the invention.
 XX
 Query Match 98.5%; Score 1133; DB 22; Length 222;
 Best Local Similarity 99.1%; Pred No. 7, 1e-118;
 Matches 220; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 TCTPGAVATRDGEAPERSPPCSYDLTGKVMLLGDTGVGKTCFLIQKDGAFSLGTFTA 61
 Db 1 TCTPGAVATRDGEAPERSPPCSYDLTGKVMLLGDTGVGKTCFLIQKDGAFSLGTFTA 60
 QY 62 TVGIDFRNKVVTVDGVVRKQLQWDAGQFRFSVTHAYYRDAQAALLLLYDITNKSSFDNI 121
 Db 61 TVGIDFXNKVVTVDGVVRKQLQWDAGQFRFSVTHAYYRDAQAALLLLYDITNKSSFDNI 120
 QY 122 RAWLTEIHEYAQRDYYVIMLGNKADMSRVRIRSEDETALAREYGVVPFLETSAKTMNVE 181
 Db 121 RAWLTEIHEYAQRDYYVIMLGNKADMSRVRIRSEDETALAREYGVVPFLETSAKTMNVE 180
 QY 182 LAFLAIAKELKYRAGHQADEPFSQIRDYVESQKRSSCCSFM 223
 Db 181 LAFLAIAKELKYRAGHQADEPFSQIRDYVYXQKRSSCCSFM 222
 RESULT 4
 AAG67156
 ID AAG67156 standard; Protein; 191 AA.
 XX
 AC AAG67156;
 XX
 DT 13-NOV-2001 (first entry)
 XX

DE Amino acid sequence of human 32712 G-protein.
 XX Human; G-protein; 32705; 23224; 27423; 32700; 32712; lung disorder;
 KW congenital anomaly; pulmonary congestion; oedema; haemorrhage;
 KW adult respiratory distress syndrome; Goodpasture's syndrome;
 KW chronic obstructive pulmonary disease; asthma; pulmonary hypertension;
 KW liver disorder; hepatic injury; jaundice; cholestasis; viral hepatitis;
 KW cirrhosis; Wilson's disease; autoimmune hepatitis; hepatic failure;
 KW brain disorder; hypoxia; cerebral ischemia; intracranial haemorrhage;
 KW acute meningitis; Parkinson's disease; Alzheimer's disease; glioma;
 KW chronic bacterial meningococcalitis; multiple sclerosis;
 KW amyotrophic lateral sclerosis; stroke; Huntington's disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200164887-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 27-FEB-2001; 2001WO-US06292.
 XX
 PR 29-FEB-2000; 2000US-0185606.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 PI Meyers RA;
 XX
 XX WPI: 2001-550182/61.
 DR N-PSDB; AAH75184.
 XX
 PT Novel human small G-protein polypeptides and polynucleotides for
 PT treating lung disorders, liver disorders and brain disorders
 XX
 PS Claim 8; Fig 26; 151pp; English.
 XX
 CC The present sequence represents a human G-protein. The specification
 CC describes 32705, 23224, 27423, 32700 or 32712 small G-proteins. The
 CC G-protein polypeptides and polynucleotides are useful as a target for
 CC diagnosis and treatment of G-protein mediated or related disorders,
 CC and for identifying agonists and antagonists for diagnosis and
 CC treatment. They are useful for treating disorders of lung (e.g.
 CC congenital anomalies, pulmonary congestion, oedema, adult respiratory
 CC distress syndrome, haemorrhage, chronic obstructive pulmonary disease,
 CC asthma, Goodpasture's syndrome and pulmonary hypertension), liver
 CC (e.g. hepatic injury, jaundice, cholestasis, viral hepatitis, cirrhosis,
 CC Wilson's disease, autoimmune hepatitis and hepatic failure), and
 CC brain (e.g. hypoxia, cerebral ischemia, intracranial haemorrhage, acute
 CC meningitis, Parkinson's disease, Alzheimer's disease, gliomas, chronic
 CC bacterial meningococcalitis, multiple sclerosis, amyotrophic lateral
 CC sclerosis, stroke and Huntington's disease).
 XX
 SQ Sequence 191 AA;

Query Match 85.0%; Score 977; DB 22; Length 191;
 Best Local Similarity 100.0%; Pred. No. 1.5e-100;
 Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 33 MLGDTGVGKTCFLIQKDGAFSLGTFTAIVGIDFRNKVVTVDGVVRKQLQWDAGQFRF 92
 Db 1 MLGDTGVGKTCFLIQKDGAFSLGTFTAIVGIDFRNKVVTVDGVVRKQLQWDAGQFRF 60
 QY 93 RSVTHAYYRDAQAALLLLYDITNKSSFDNIRAWLTEIHEYAQRDYYVIMLGNKADMSERV 152
 Db 61 RSVTHAYYRDAQAALLLLYDITNKSSFDNIRAWLTEIHEYAQRDYYVIMLGNKADMSERV 120
 QY 153 IRSEDETALAREYGVVPFLETSAKTMNVELAFLAIAKELKYRAGHQADEPFSQIRDYVES 212
 Db 121 IRSEDETALAREYGVVPFLETSAKTMNVELAFLAIAKELKYRAGHQADEPFSQIRDYVES 180
 QY 213 QKRSSCCSFM 223
 Db 181 QKRSSCCSFM 191

RESULT 5
AAM39910
ID AAM39910 standard; Protein; 191 AA.
XX AC AAM39910;
XX DT 22-OCT-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 3055.
XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX OS Homo sapiens.
XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US34263.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-052317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
XX N-PSDB; AAI59066.
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX Example 4; SEQ ID NO 3055; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX assays for receptor activity, cancer diagnosis and therapy, drug screening,
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.
XX Seqence 191 AA;
XX Query Match 85.0%; Score 977; DB 22; Length 191;
XX Best Local Similarity 100.0%; Pred. No. 1.5e-100;
XX Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX 33 MLLGDTGVGKTCFLIOFKDGAFLSGTFFIATVGIDFRNKVTVDSGVRVKLQINDTAGQERF 92

Db 1 MLLGDTGVGKTCFLIOFKDGAFLSGTFFIATVGIDFRNKVTVDSGVRVKLQINDTAGQERF 60
QY 93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTHEIHYAQRDVTMLLGNKADMSERV 152
Db 61 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTHEIHYAQRDVTMLLGNKADMSERV 120
QY 153 IRSEGETLAREYGVFPFLETSARTGMNVELAFIAIAKELKYRAGHQADEPSPFOIRDYVES 212
Db 121 IRSEGETLAREYGVFPFLETSARTGMNVELAFIAIAKELKYRAGHQADEPSPFOIRDYVES 180
QY 213 QKKRSSCCSFM 223
Db 181 QKKRSSCCSFM 191
RESULT 6
ABE76426
ID ABB76426 standard; Protein; 191 AA.
XX AC ABB76426;
XX DT 27-AUG-2002 (first entry)
XX DE Human Ras protein 3 (RASP-3).
XX KW Ras protein 3; RASP-3; human; cancer; immune disease; cytostatic;
KW immunosuppressive; antiinflammatory; signal transduction.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Region 1..19 /note= "transforming protein P21 family signature"
XX FT Region 4..11 /note= "ATP/GTP-binding site, motif A (P-loop)"
XX FT Region 22..38 /note= "transforming protein P21 family signature"
XX FT Region 40..62 /note= "transforming protein P21 family signature"
XX FT Region 63..80 /note= "Ran family signature sequence"
XX FT Region 102..115 /note= "transforming protein P21 family signature"
XX FT Region 130..168 /note= "Ran family signature sequence"
XX FT Region 137..159 /note= "transforming protein P21 family signature"
XX US6391580-B1.
XX 21-MAY-2002.
XX 08-MAY-1998; 98US-0075454.
XX 12-DEC-1996; 96US-0766551.
XX (INCY-) INCYTE GENOMICS INC.
XX Hillman JL, Tang YT, Lal P, Guegler KJ, Corley NC, Patterson C;
XX Batra S, Baughn MR;
XX WPI; 2002-498774/53.
XX N-PSDB; ABN83696.
XX New isolated polynucleotides encoding Ras proteins designated RASP-1
XX and RASP-4, for diagnosing, preventing and treating disorders
XX associated with cell proliferation, particularly cancer and immune
XX disorders -
XX Example; Column 47-48; 34pp; English.
XX The present sequence is the protein sequence of novel human Ras

CC protein 3 (RASP-3). RASP-3 shares 72% sequence identity with
CC Rab26. It includes regions that resemble the signature sequences
CC of the GTP-binding Ras superfamily, and family signatures of Ras,
CC a subfamily of Ras proteins which function in nucleocytoplasmic
CC transport, RNA synthesis, processing and export, and cell cycle
CC checkpoint control, as well as family signatures of transforming
CC protein p21, a subfamily of GTP-binding Ras proteins which have
CC been implicated in a number of tumours. Northern analysis showed
CC expression of RASP-3 in haematopoietic and immunological cDNA
CC libraries, all of which were associated with inflammation and the
CC immune response. The invention provides 7 novel human Ras proteins
CC (RASP-1 to -7) and polynucleotides, expression vectors, host cells,
CC antibodies, agonists and antagonists. It also provides methods for
CC diagnosing, treating or preventing disorders associated with RASP
CC expression, especially cancer and immune disorders.
XX
SQ Sequence 191 AA;

Query Match 85.0%; Score 977; DB 23; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.5e-100;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 MLLGDTGVGKTCFLIQFDGAFSLGTFATVIGIDFRNKVTVVDGVRVKLIQIWDTAGQERF 92
DB 1 MLLGDTGVGKTCFLIQFDGAFSLGTFATVIGIDFRNKVTVVDGVRVKLIQIWDTAGQERF 60
QY 93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLLGNKADMSSERV 152
DB 61 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLLGNKADMSSERV 120
QY 153 IRSEDETLAREYGVFPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
DB 121 IRSEDETLAREYGVFPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 180
QY 213 QKKRSSCCSFM 223
DB 181 QKKRSSCCSFM 191

RESULT 7

AAM52190
ID AAM52190 standard; Protein; 191 AA.

XX AC AAM52190;
XX DT 07-FEB-2002 (first entry)
XX DE Human degranulation regulator SEQ ID NO 3.
XX KW Degranulation; mast cell; human; mouse; antiallergic.
XX OS Homo sapiens.

XX FN WO200179478-A1.
XX PD 25-OCT-2001.
XX PF 16-APR-2001; 2001WO-JP03268.

XX PR 19-APR-2000; 2000JP-0118408.
XX PA (DAIN) DAINIPPON PHARM CO LTD.
XX PI Yamada T, Ido M;

XX DR WPI; 2002-041335/05.
XX DR N-PSDB; ABA02774.

XX PT Mast cell degranulation controller for treatment of allergies

XX PS Claim 11; Page 64-65; 85pp; Japanese.

XX CC The invention relates to a protein for regulating degranulation of mast

CC cells (degranulation regulators) and the encoding polynucleotides, with
CC antiallergic activity, used in the treatment of allergies associated
CC with degranulation of mast cells.

XX SQ Sequence 191 AA;

Query Match 84.3%; Score 969; DB 23; Length 191;
Best Local Similarity 99.0%; Pred. No. 1.1e-99;
Matches 189; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 33 MLLGDTGVGKTCFLIQFDGAFSLGTFATVIGIDFRNKVTVVDGVRVKLIQIWDTAGQERF 92
DB 1 MLLGDSGVGKTCFLIQFDGAFSLGTFATVIGIDFRNKVTVVDGVRVKLIQIWDTAGQERF 60
QY 93 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLLGNKADMSSERV 152
DB 61 RSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLLGNKADMSSERV 120
QY 153 IRSEDETLAREYGVFPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 212
DB 121 IRSEDETLAREYGVFPFLETSAKTGMNVELAFLAIAKELKYRAGHQADEPSFQIRDYVES 180
QY 213 QKKRSSCCSFM 223
DB 181 QKKRSSCCSFM 191

RESULT 8

AAM52189
ID AAM52189 standard; Protein; 191 AA.

XX AC AAM52189;
XX DT 07-FEB-2002 (first entry)
XX DE Mouse degranulation regulator SEQ ID NO 2.
XX KW Degranulation; mast cell; human; mouse; antiallergic.
XX OS Mus musculus.

XX FH Key Location/Qualifiers
XX FT Misc-difference 83
XX FT /note= "Encoded by CAG"

XX PN WO200179478-A1.
XX PD 25-OCT-2001.
XX PF 16-APR-2001; 2001WO-JP03268.

XX PR 19-APR-2000; 2000JP-0118408.

XX PA (DAIN) DAINIPPON PHARM CO LTD.

XX PI Yamada T, Ido M;

XX DR WPI; 2002-041335/05.
XX DR N-PSDB; ABA02773.

XX PT Mast cell degranulation controller for treatment of allergies

XX PS Claim 10; Page 63-64; 85pp; Japanese.

XX CC The invention relates to a protein for regulating degranulation of mast
XX cells (degranulation regulators) and the encoding polynucleotides, with
XX antiallergic activity, used in the treatment of allergies associated
XX with degranulation of mast cells.

XX SQ Sequence 191 AA;

Query Match 82.3%; Score 946; DB 23; Length 191;
Best Local Similarity 96.3%; Pred. No. 4.2e-97;

```
Matches 184; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 33 MLLGDTGVGKTCFLIQKDGAFLSGTFIATVGIDFRNKVVTVVGVKVKLIQWDTAGQERF 92
DB 1 MLLGDSGVGKTCFLIQKDGAFLSGTFIATVGIDFRNKVVTVVGVKVKLIQWDTAGQERF 60
QY 93 RSVTHAYYRDAQAALLLLYDITNKSSFDNIRAWLTEIHEYAQRDQVIMLLGNKADMSERV 152
DB 61 RSVTHAYYRDAQAALLLLYDITNKSSFDNIRAWLTEIHEYAQRDQVIMLLGNKADMSERV 120
QY 153 IRSEDEGTILAREGVVPFLETSKTMNVELAFIAIAKELKYRAGHQADSPSFQIRDYVES 212
DB 121 IRSEDEGTILAREGVVPFLETSKTMNVELAFIAIAKELKYRAGHQADSPSFQIRDYVES 180
QY 213 QKKRSSCCSFM 223
DB 181 QKKRSSCCSFM 191

RESULT 9
AAM521188
ID AAM521188 standard; protein; 191 AA.
XX
AC AAM521188;
XX
DT 07-FEB-2002 (first entry)
XX
DE Degranulation regulator SEQ ID NO 1.
XX
KW Degranulation; mast cell; human; mouse; antiallergic.
XX
OS Homo sapiens.
XX
Mus musculus.
XX
FH Key Location/Qualifiers
FT Misc-difference 45 /label= Ala, Val
FT Misc-difference 115 /label= Val, Met
FT Misc-difference 138 /label= Met, Leu
FT Misc-difference 165 /label= Arg, His
FT Misc-difference 167 /label= Pro, Ala
XX
WO200179478-A1.
XX
PN
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-JP03268.
XX
PR 19-APR-2000; 2000JP-0118408.
XX
PA (DAIN ) DAINIPPON PHARM CO LTD.
XX
PI Yamada T, Ido M;
XX
WPI; 2002-041335/05.
XX
PT Mast cell degranulation controller for treatment of allergies -
XX
PS Claim 1; Page 63; 85pp; Japanese.
XX
The invention relates to a protein for regulating degranulation of mast
CC cells (degranulation regulators) and the encoding polynucleotides, with
CC antiallergic activity, used in the treatment of allergies associated
CC with degranulation of mast cells.
XX
Sequence 191 AA;
Query Match 81.7%; Score 940; DB 23; Length 191;
Best Local Similarity 96.3%; Pred. No. 2e-96;
```

```
Matches 184; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 33 MLLGDTGVGKTCFLIQKDGAFLSGTFIATVGIDFRNKVVTVVGVKVKLIQWDTAGQERF 92
DB 1 MLLGDSGVGKTCFLIQKDGAFLSGTFIATVGIDFRNKVVTVVGVKVKLIQWDTAGQERF 60
QY 93 RSVTHAYYRDAQAALLLLYDITNKSSFDNIRAWLTEIHEYAQRDQVIMLLGNKADMSERV 152
DB 61 RSVTHAYYRDAQAALLLLYDITNKSSFDNIRAWLTEIHEYAQRDQVIMLLGNKADMSERV 120
QY 153 IRSEDEGTILAREGVVPFLETSKTMNVELAFIAIAKELKYRAGHQADSPSFQIRDYVES 212
DB 121 IRSEDEGTILAREGVVPFLETSKTMNVELAFIAIAKELKYRAGHQADSPSFQIRDYVES 180
QY 213 QKKRSSCCSFM 223
DB 181 QKKRSSCCSFM 191

RESULT 10
AAY30132
ID AAY30132 standard; Protein; 191 AA.
XX
AC AAY30132;
XX
DT 27-OCT-1999 (first entry)
XX
DE An exocytotic protein designated Exo2.
XX
KW Exocytotic protein; Exo2; exocytosis; Chediak-Higashi Syndrome;
XX inflammation.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Misc-difference 5 /note= "encoded by GAN"
FT Misc-difference 27 /note= "encoded by TTN"
XX
PN WO9942586-A2.
XX
PD 26-AUG-1999.
XX
PF 23-FEB-1999; 99WO-US03909.
XX
PR 26-MAY-1998; 98US-0086650.
XX
PR 23-FEB-1998; 98US-0075534.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Anderson D, Fisher J, Huang B, Lorens J, Luo Y;
XX Shen M;
XX
WPI; 1999-518605/43.
XX
N-PSDB; AAX86720.
XX
PT New exocytotic proteins useful for diagnosis and treatment of
XX exocytosis-mediated conditions and in drug screening
XX
PS Claim 12; Fig 7; 53pp; English.
XX
The present sequence represents an exocytotic protein designated Exo2.
CC The protein has an inhibitory effect on exocytosis (i.e. the fusion of
CC secretory vesicles with the cellular plasma membrane) and can be
CC administered therapeutically to treat or prevent exocytosis-mediated
CC disorders e.g. Chediak-Higashi Syndrome (CHS) or conditions involving
CC undesirable release of compounds via exocytosis e.g. inflammation
CC mediated by the release compounds such as histamine. Exo2 is useful
CC diagnostically and to produce antibodies useful to purify the proteins
CC and therapeutically to reduce or eliminate the biological activity of
CC the protein. Exo2 nucleic acids can be used therapeutically to increase
CC Exo2 activity in cells by known gene therapy techniques. They can also
```

CC be used to produce probes or primers to isolate Exo2 proteins from other
CC organisms, especially humans. The nucleic acids, host cells and proteins
CC are useful in screening assays to identify binding agents, especially
CC drug screening assays to identify agonists and antagonists useful
CC therapeutically to enhance or reduce Exo2 activity.
XX
SQ Sequence 191 AA;

Query Match 80.5%; Score 926; DB 20; Length 191;
Best Local Similarity 94.8%; Pred. No. 7.2e-95;
Matches 181; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 33 MLGGDTGKTCFLIQKDGAFLSGTFIATVGIDFRNKVTVVGVKVLQIWDTAGQERF 92
Db 1 MLGGSGVGKTCFLIQKDGAFLSGTFIATVGIDFRNKVTVVGVKVLQIWDTAGQERF 60
QY 93 RSVTHAYYRDAQALLLLYDITNKSFDNIRAWLITEIHEYAQRDVVIMLGNKADMSERV 152
Db 61 RSVTHAYYRDAQALLLLYDITNKSFDNIRAWLITEIHEYAQRDVVIMLGNKADMSERV 120
QY 153 IRSEGETIAREYGVFPFLETSKAGTMNVELAFATAKELKYRAGHQADPEFSQIRDYVES 212
Db 121 IRSEGETIAREYGVFPFLETSKAGTMNVELAFATAKELKYRAGHQADPEFSQIRDYVES 180
QY 213 QKKRSCCSFM 223
Db 181 QKKRSCCSFV 191

RESULT 11
ABG17244
ID ABG17244 standard; Protein; 158 AA.

XX AC ABG17244;

DT 18-FEB-2002 (first entry)
DE Novel human diagnostic protein #17235.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS81431.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

XX Claim 20; SEQ ID NO 47603; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 158 AA;

Query Match 67.0%; Score 770; DB 22; Length 158;
Best Local Similarity 99.3%; Pred. No. 1.4e-77;
Matches 148; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTGTPCAVATRDGEAPERSPPCPSYDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTFI 60
Db 10 MTGTPCAVATRDGEAPERSPPCPSYDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTFI 69
QY 61 ATVGIDFRNKVTVVGVKVLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSFDN 120
Db 70 ATVGIDFRNKVTVVGVKVLQIWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSFDN 129
QY 121 IRAWLITEIHEYAQRDVVIMLGNKADMS 149
Db 130 IRAWLITEIHEYAQRDVVIMLGNKADMS 158

RESULT 12
ABB71578
ID ABB71578 standard; Protein; 666 AA.

XX AC ABB71578;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 41526.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL15681.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX Disclosure; SEQ ID NO 41526; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 666 AA;

Query Match 60.6%; Score 697; DB 22; Length 666;

Best Local Similarity 59.8%; Pred. No. 1.6e-68;

Matches 137; Conservative 33; Mismatches 49; Indels 10; Gaps 3;

QY 3 GTPCAVAT---RDGEA-----PERSPPCPSVDLTKGKWLMDGTGKTCFLIQKDGKA 53

DB 433 GSEGASATLCKNAGRALIRMISSKAPKEEDEFKGVIMLGDGSGVKTSLIRFRDGR 492

QY 54 FLSGTITATVGIDFRNKNVTVGVRVKLQIWDTAGOERFSVTHAYYRDAQALLLYDIT 113

DB 493 YVPSYFLSTVGIDFRNKNVVVDGTRVKLQIWDTAGOERFSVTHAYYRDAHALLLYDVT 552

QY 114 NKSSFNIRAWLFEIHEYAQRDVVIMLLGNKADMS-SERVIRSEGETTLAREYGVPPLET 172

DB 553 NKTTYDNIRAWLGEIREYAQEDVVVLIGNKADCSGSRQVKREDGERLGRHNVPFMEET 612

QY 173 SAKTGMNVELAFIAIAKELKYRAGHQADEPFSFOIRDYVVSQKRSKCS 221

DB 613 SAKTGLNVELSFATAVRAQLKSRGHEHDDGKFNVDHFRDNTKARSVCA 661

RESULT 13

AAW80747
ID AAW80747 standard; Protein; 190 AA.

AC AAW80747;

DT 17-DEC-1998 (first entry)

DE Human RAB protein, SRAB.

XX Human; RAB protein; SRAB; vesicle trafficking disorder; epilepsy; cancer;
KW Huntington's disease; Parkinson's disease; schizophrenia; choroideraemia;
KW viral infection; therapy; autoimmune sialosis; cystic fibrosis;
KW diabetes mellitus; hyperglycaemia; tubulointerstitial nephritis;
KW hypoglycaemia; pancreatic enzyme deficiency; pancreatitis;
KW polycystic renal disease.

OS Homo sapiens.

XX WO9842839-A1.

XX 01-OCT-1998.

XX 25-MAR-1998; 98WO-US05871.

XX 26-MAR-1997; 97US-0824873.

XX (INCY-) INCYTE PHARM INC.

XX Guegler KJ, Hillman JL;

XX WPI; 1998-542282/46.

XX N-PSDB; AAV65197.

XX New human Rab protein, SRAB - useful for treating disorders
PT associated with SRAB expression, including vesicle trafficking,
PT viral infection, and cancer

XX Claim 1; Fig 1; 31pp; English.

XX

CC This sequence is the human RAB protein, SRAB, of the invention. The
CC SRAB protein can be used in compositions useful in the diagnosis,
CC prevention, or treatment of disorders associated with vesicle trafficking
CC (including epilepsy, Huntington's disease, Parkinson's disease and
CC schizophrenia), cancer, or viral infection. Host cells containing the DNA
CC sequence can be used to produce SRAB recombinantly. The viral infections
CC that can be treated using SRAB include those caused by retroviruses
CC particularly HIV and HTLV, hepadnaviruses, particularly hepatitis C,
CC hantaviruses, herpesviruses, and arboviruses. The cancers that can be
CC treated include adenocarcinoma, leukaemia, lymphoma, melanoma and
CC sarcoma, particularly cancers of the endocrine, gastrointestinal and
CC nervous systems and cancers of the adrenal gland, brain, breast, colon,
CC oesophagus, kidney, liver, lung, ovaries, pancreas, pituitary gland,
CC prostate, salivary gland, stomach, thyroid, and uterus. The protein can
CC also be used to treat conditions such as autoimmune sialosis,
CC choroideaemia, cystic fibrosis, diabetes mellitus, hyperglycaemia,
CC hypoglycaemia, tubulointerstitial nephritis, pancreatic enzyme
CC deficiency, pancreatitis, and polycystic renal disease.

XX Sequence 190 AA;

Query Match 57.9%; Score 666; DB 19; Length 190;

Best Local Similarity 67.0%; Pred. No. 7.4e-66;

Matches 126; Conservative 25; Mismatches 37; Indels 0; Gaps 0;

QY 33 MLLGDTGVGKTCFLIQKDGAFLSGTFIATVGIDFRNKNVTVGVRVKLQIWDTAGOERF 92

DB 1 MLVGDSGVGKTCFLIQKDGAFLSGTFIATVGIDFRNKNVTVGVRVKLQIWDTAGOERF 60

QY 93 RSVTHAYYRDAHALLLLYDVTNKSSFNIRAWLFEIHEYAQRDVVIMLLGNKADMSERV 152

DB 61 RSVTHAYYRDAHALLLLYDVTNKASFDNIQAWLFEIHEYAQRDVVIMLLGNKADMSERV 120

QY 153 IRSEGETTLAREYGVPPLETSAKTGMNVELAFIAIAKELKYRAGHQADEPFSFOIRDYVES 212

DB 121 VKREDGEKLAKEYGLPFPMTSAKTGLNVDIAFTAIKELQKRSKAPSEPRFRLHDYVKR 180

QY 213 QKKRSKCC 220

DB 181 EGRGASCC 188

RESULT 14

AAU17123
ID AAU17123 standard; Protein; 139 AA.

XX AAU17123;

XX 07-NOV-2001 (first entry)

XX Novel signal transduction pathway protein, Seq ID 688.

XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell-anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS;
KW acquired immune deficiency syndrome.

XX Homo sapiens.

XX WO200154733-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01312.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244674.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465460/50.
XX N-PSDB; AAS27040.
XX
XX Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders

```
PS Claim 1; SEQ ID No 688; 880pp; English.
XX
XX The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAU17059-AAU1783 represent novel signal transduction
CC pathway protein, amino acid sequences of the invention.
XX

Query Match 46.5%; Score 535; DB 22; Length 139;
Best Local Similarity 71.5%; Pred. No. 1.9e-51;
Matches 98; Conservative 19; Mismatches 20; Indels 0; Gaps 0;

QY 84 WDTAGQFRFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLLGN 143
DB 1 WDTAGQFRFRSVTHAYYRDAHALLLYDITNKASFDNIQAWLTIHEYAQRDVHVALMLLGN 60
QY 144 KADMSSSERVIRSEDETLAREYGVFFLETSAKGMVVELAFIAKELKTRAGHQADEPS 203
DB 61 KVDSAHERVVKREDGEKLAKEYGLPFMETSAKTGLNVDLAFIAKELKORSKMFSEPR 120
QY 204 FQIRDYVESOKRRSSC 220
DB 121 FRLHDYVKREGASCC 137

RESULT 15
ABP41333
ID ABP41333 standard; Protein; 221 AA.
XX
XX ABP41333;
XX
XX 23-AUG-2002 (first entry)
XX
XX Human ovarian antigen HCGMA67, SEQ ID NO:2465.
XX
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX ovarian cancer; breast cancer; reproductive system disorder;
XX infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
XX PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
XX inflammatory condition; immune disorder; blood disorder;
XX cardiovascular disorder; respiratory disorder; neurological disorder;
XX gastrointestinal disorder; urinary system disorder; drug screening;
XX gene therapy; chromosome mapping; forensic analysis;
XX antibody preparation; cytostatic; immunomodulatory; neuroprotective;
XX antiinflammatory; gynaecological; reproductive.
XX
XX Homo sapiens.
XX
XX OS
XX
XX WO200200677-A1.
XX
XX
XX 03-JAN-2002.
XX
XX
XX 07-JUN-2001; 2001WO-US18569.
XX
```

```
PR 07-JUN-2000; 2000US-209467P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX PA
XX PI Birse CE, Rosen CA;
XX
XX WPI; 2002-147878/19.
XX N-PSDB; ABQ54410.
XX
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
XX useful in the prevention, treatment and diagnosis of cancer (e.g.
XX ovarian cancer), immune disorders, cardiovascular disorders and
XX neurological diseases -
XX
XX Claim 11; SEQ ID No 2465; 2922pp; English.
XX
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX encompasses polypeptides 90% identical and polynucleotides 95% identical
XX to the sequences of the invention. The invention additionally relates to
XX recombinant vectors and host cells comprising human ovarian antigen
XX polynucleotides, antibodies against human ovarian antigens, and the use
XX of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX treating, prognosing or preventing various ovary and/or breast-related
XX disorders. Such conditions include ovarian cancer and breast cancer, and
XX metastatic tumours of ovarian or breast origin, reproductive system
XX disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
XX shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
XX vaginitis), immune disorders (e.g., congenital and acquired
XX immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
XX blood-related disorders (e.g., anaemia), cardiovascular disorders,
XX respiratory disorders, neurological disorders, gastrointestinal disorders
XX and urinary system disorders. Ovarian antigen polypeptides and
XX polynucleotides may also be used in screening for compounds which
XX modulate ovarian antigen expression or activity. The polynucleotides may
XX further be used for gene therapy, chromosome mapping, in the
XX identification of individuals and in forensic analysis, and the
XX polypeptides may be used as food additives or to prepare antibodies
XX useful in disease diagnosis, drug targeting and phenotyping. The present
XX sequence represents a human ovarian antigen of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 221 AA;
```

```
Query Match 44.7%; Score 514.5; DB 23; Length 221;
Best Local Similarity 48.3%; Pred. No. 7.4e-49;
Matches 100; Conservative 43; Mismatches 59; Indels 5; Gaps 2;

QY 16 PERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQKDCAFLSGTFIATGVIDRKNKVVTV 75
DB 9 PGSTHASAKTYDYFLKLLIGDSGVGKTCVLFRESDAF-NSTFIISTIGDKIRFIELD 67
QY 76 GYRVKLIQIWDTAGQFRFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRD 135
DB 68 GKRIKLIQIWDTAGQFRFRITAYYRGAMGIMLVYDITNEKSFDRNIRNIRIEHASAD 127
QY 136 VYIMLGNKADMSSSERVIRSEDETLAREYGVFFLETSAKGMVVELAFIAKELKYRA 195
DB 128 VEKMLGNKCDVNDKRQVSKERGEKALDYGIKFMETSAKANINVENAFFTLARDIKAKM 187
QY 196 GHQAD----EPSFQIRDYVESOKRRSS 218
DB 188 DKKLEGNSPQGSNQGVKITPPDQOKRSS 214
```

Search completed: November 17, 2002, 16:39:23
Job time : 68 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 17, 2002, 16:41:03 ; Search time 2651 Seconds
(without alignments)
2448.105 Million cell updates/sec

Title: US-09-817-199A-2
Perfect score: 1150
Sequence: 1 MTCTPGAVATRDGEAPERSP.....FQIRDYVESQKRSSCCSPM 223

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=framet_p2n_model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09817199/runat_13112002_134902_5165/app.query.fasta_1.391
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09817199 @CGN_1.1.1616 @runat_13112002_134902_5165 -NCPU=6 -ICPU=3
-NO_XLPYX -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: gb_wa.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*

- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1150	100.0	1700	9	BC016615 Homo sapi
2	1150	100.0	1977	9	AK098068 Homo sapi
3	1145	99.6	1116	6	AX236082 Sequence
4	1081	94.0	690	10	AF233582 Mus muscu
5	977	85.0	576	6	AX236084 Sequence
6	780.5	67.9	1513	9	BC007681 Homo sapi
7	757	65.8	1320	9	AB027137 Homo sapi
8	754	65.6	1098	10	U18771 Rattus norv
9	753	65.5	573	9	AF498952 Homo sapi
10	731	63.6	1340	6	AK062279 Sequence
11	710.5	61.8	3105	9	AK054846 Homo sapi
12	649.5	56.5	2536	3	AY061826 Drosophil
13	521.5	45.3	1337	10	BC019990 Mus muscu
14	518.5	45.1	1074	8	AB024994 Cicer ari
15	517.5	45.0	2145	3	DDISAS1A M34456 D.discoidu
16	517.5	45.0	10034	2	AC115606 Dictyoste
17	516	44.9	638	10	S53270 MEL-RAS-rel
18	516	44.9	765	5	DYGORAZ2 S33270 MEL-RAS-rel
19	516	44.9	911	8	AY084345 Arabidops
20	516	44.9	2048	9	BC002977 Homo sapi
21	516	44.9	2818	9	AK025165 Homo sapi
22	513.5	44.7	624	9	AF498943 Homo sapi
23	513.5	44.7	660	9	HSRAB8 X56741 H.sapiens m
24	513.5	44.7	740	10	RNU53475 Rattus norv
25	513.5	44.7	760	4	CFRAB8 X56385 Canine rab8
26	513.5	44.7	782	8	PEAGTPBP09 D12548 Pea mRNA fo
27	513.5	44.7	1265	9	AB038995 Homo sapi
28	512.5	44.6	1080	8	BVRAB1 Z49152 B.vulgaris
29	512	44.5	1980	9	S53268 Homo sapien
30	511.5	44.5	2497	6	AX285074 Sequence
31	511.5	44.5	2497	6	AX285080 Sequence
32	511.5	44.5	2497	6	AX285089 Sequence
33	511.5	44.5	2497	6	AX285091 Sequence
34	510	44.3	919	8	DCJ001367 AJ001367 Daucus ca
35	509.5	44.3	1082	8	LJRA8D J73947 L.japonicus
36	509	44.3	955	8	S66160 r1c1ar-as-re
37	509	44.3	1163	8	LJRA8A J73944 L.japonicus
38	508.5	44.2	1613	3	AB006189 Drosophil
39	508.5	44.2	2349	3	AY060425 AY060425 Drosophil
40	507.5	44.1	979	8	LJRA1C J73932 L.japonicus
41	507	44.1	681	8	AY056351 Arabidops
42	507	44.1	821	8	AY052333 Arabidops
43	507	44.1	986	8	AB079023 Nicotiana
44	507	44.1	1025	8	AY086756 Arabidops
45	507	44.1	1033	8	AB079021 Nicotiana

ALIGNMENTS

RESULT 1

```

BC016615
LOCUS      BC016615      1700 bp      mRNA      linear      PRI 05-NOV-2001
DEFINITION Homo sapiens, similar to RAB37, member of RAS oncogene family,
clone MGC:21391 IMAGE:4520191, mRNA, complete cds.
ACCESSION  BC016615
VERSION     BC016615.1 GI:16741620
KEYWORDS   MGC.
SOURCE      Homo sapiens.
ORGANISM    Homo sapiens.
REFERENCE   1 (bases 1 to 1700)
AUTHORS     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE       Strausberg, R.
JOURNAL     Direct Submission
SUBMITTED   31-OCT-2001 National Institutes of Health, Mammalian
INSTITUTE   Gene Collection (MGC), Cancer Genomics Office, National Cancer
INSTITUTE   Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
INSTITUTE   USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
            Email: cgapbs@mail.nih.gov
            Tissue procurement: DCTD/OT
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E.-E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: villalona@bcm.tmc.edu.
            Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
            A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
            Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E.-E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 28 Row: J Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
Location/Qualifiers
1..1700
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:21391 IMAGE:4520191"
/tissue_type="Prostate, adenocarcinoma."
/clone_lib="NIH_MGC_91"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
3..674
/codon_start=1
/product="Similar to RAB37, member of RAS oncogene family"
/protein_id="AAH16615.1"
/db_xref="GI:16741621"
/translation="MTGPGAVATRDGEAPERSPPSPSYDLTGKVMILGDTGVGKTC
FLIKDGAFLSCFIATVGIIDFRNKVTVVDGVKVLQIWDTAQCEFRSFWTHAYRD
AQALLLDYINKSFENIRAWLIEIHEYAQRDVVIMLLGNKADMSSEVIRSDGET
LRGEVFPFLETSKATGMNVELAFALAKELKIRAGHQADPEFSQIRDYVESQKRSS
CCSPM"

BASE COUNT 437 a 444 c 460 g 359 t
ORIGIN
Alignment Scores:
Pred. No.: 1.77e-108 Length: 1700
Score: 1150.00 Matches: 223
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-817-199A-2 (1-223) x BC016615 (1-1700)

Qy 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
|||||

```

```

Db 3 ATGACGGGCACGCCAGCGCCGCTTGCCACCGGGATGGCGAGGCCCGCCCGAGCGCTCCCG 62
Qy 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuGlyAspThrGlyVal 40
|||||
Db 63 CCTCGAGTCCGAGCTACGACCTCAGGGCAAGGTGATGCTTCTGGGAGACACAGCGCTC 122
Qy 41 GlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheile 60
|||||
Db 123 GCGAAACATGTTTCTGATCAATTCAAAGACGGGGCTTCTGTCGCGAACCCTTCAT 182
Qy 61 AlaThrValGlyLeuAspPheArgLysValValThrValAspGlyValArgValLys 80
|||||
Db 183 GCCACCGTCGGCATAGACTTCAGGAACAAGTGGTGAATGCTGATGCGCTGAGAGTGAAG 242
Qy 81 LeuGlnIleTyrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
|||||
Db 243 CTGCGATCTGGGACACCGCTGGGAGGAACGGTTCCGAAGCGTCAACCATGCTTTATTAC 302
Qy 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
|||||
Db 303 AGAGATGCTCAGCGCTTGCTTCTGCTGATGACATCACCAACAATCTTCTTCGACAAC 362
Qy 121 IleArgAlaThrLeuThrGluLeuHisGluTyrAlaGlnArgAspValValIleMetLeu 140
|||||
Db 363 ATCAGGGCGCTGGCTCACTGAGATTGATGATGATGATGATGATGATGATGATGATG 422
Qy 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
|||||
Db 423 CTAGCAACAAGCGGATGATGACAGCGAAGAGATGATCCGTTCCGACAGGAGAGACC 482
Qy 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
|||||
Db 483 TTGGCCAGGGAGTACGGTGTCTCCCTGCGAGACGAGCGCCCAAGACTGCCATGAATGTG 542
Qy 181 GluLeuAlaPheLeuAlaLeuLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
|||||
Db 543 GAGTAGCCCTTCTGCCATCGCCCAAGGAACATAACCGGGCGGGCATCAGCGGAT 602
Qy 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
|||||
Db 603 GAGCCAGCTTCCAGATCCGAGACTATGTAGATGCCAGAGAGCGCTCCAGCTGCTGC 662
Qy 221 SerPheMet 223
|||||
Db 663 TCCTTCATG 671

RESULT 2
AK098068
LOCUS      AK098068      1977 bp      mRNA      linear      PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ40749 fis, clone TRACH2000540, highly similar
to Mus musculus GTPase Rab37 (Rab37) mRNA.
ACCESSION  AK098068
VERSION     AK098068.1 GI:21758001
KEYWORDS   oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens trachea cDNA to mRNA, clone lib:TRACH2
            clone:TRACH2000540.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Ninomiya, K., Wagatsuna, M., Kanda, K., Kondo, H., Yokoi, T.,
            Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,
            Kamiyama, K., Katsuta, N., Sato, K., Tanikawa, M., Yanazaki, M.,
            Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakamatsu, A.,
            Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K.,
            Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y.,
            Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K.,
            Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,
            Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.
            NEDO human cDNA sequencing project
            Unpublished
            2 (bases 1 to 1977)
            Isogai, T. and Yamamoto, J.

```

TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology (RAB); cDNA library construction; Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing; HRI and RAB; annotation: HRI and RAB.

FEATURES

source

1..1977
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="TRACH2000540"
 /tissue_type="trachea"
 /clone_lib="TRACH2"
 /note="cloning vector: pWE18SFL3"

CDS
 57..728
 /note="unnamed protein product"

/codon_start=1
 /protein_id="BAC05227.1"
 /db_xref="GI:21758002"
 /translation="MTGTPGAVATRDGEAPSPSPSYDLTGKVMLLGDTGVGKTC FLIOFKDGAFLSGTFLATVGDIFRNKVVTVGVKQIWDTAGOERFSVTHAYRD AQLLLYDINXKSFONIRAWLEIHEYAQRDVVIMLLGNKADMSSEVRVSEDET LAREYGVFLETSKTKMNNVELAFALAKELKYRAGHQDEPFOIRDYVESQKRSS CCSPM"
 BASE COUNT 442 a 560 c 539 g 436 t

Alignment Scores:

Pred. No.: 2,16e-108 Length: 1977
 Score: 1150.00 Matches: 223
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-817-199A-2 (1-223) x AK098068 (1-1977)

QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
 |||||
 DB 57 ATGACGGGCACGCCAGGCCGCTGCCACCCGGATGGGAGGCCGCCGAGCGCTCCCG 116
 |||||
 QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
 |||||
 DB 117 CCTGTCAGTCCGAGCTACGACCTCACGGCAAGGTGATGCTTCTGGGAGACACAGCGCTC 176
 |||||
 QY 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
 |||||
 DB 177 GCGAAACATGTTCTGTATCCCAATTCAAAGACGGGGCTTCTGTCCGGAACCTTCATP 236
 |||||
 QY 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
 |||||
 DB 237 GCCACCGTCGGCATAGACTTCAGAACAGGTGGTGACTGTGGATGGCTGAGAGTGAG 296
 |||||
 QY 81 LeuGlnIleTyrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
 |||||
 DB 297 CTGCAGATCTGGGACACCGCTGGCAGGAACGGTTCGGAAGGCTACCCATGCTATTAT 356
 |||||
 QY 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerPheAspAsn 120
 |||||
 DB 357 AGATGCTCAGGCCCTGCTTCTGTGTATGACATCACCAAAATCTTCTTTCGACAC 416
 |||||
 QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
 |||||
 DB 417 ATCAGGCCCTGCTCAGTATGATGATGCCAGAGGACGCTGGTGATGCTGCTG 476
 |||||
 QY 141 LeuGlyAsnLysAlaAspMetSerGluArgValIleArgSerGluAspGlyGluThr 160
 |||||

DB 477 CTAGCAACAAGCGGATATGAGCAGCAAGAGAGTATCCGTCCGAAGCAGGAGACC 536
 QY 161 LeuAlaAAGGlyTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
 |||||
 DB 537 TTGGCCAGGGAGTAGGGTTCCTTCTGGAGACAGCCCAAGACTGGCATGAATGTG 596
 |||||
 QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
 |||||
 DB 597 GAGTTAGCTTCTGGCCATCCCAAGGAACCTGAAATACCGCGCGGCATCAGCGGAT 656
 |||||
 QY 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerCysCys 220
 |||||
 DB 657 GAGCCAGCTCCAGATCCGAGACTATGAGTCCCAAGAACGCTCCAGCTGCTGC 716
 |||||
 QY 221 SerPheMet 223
 |||||
 DB 717 TCCTTCAAG 725
 |||||
 RESULT 3
 AX236082
 LOCUS AX236082 1116 bp DNA linear PAT 26-SEP-2001
 DEFINITION Sequence 13 from Patent WO0164887.
 ACCESSION AX236082
 VERSION AX236082.1 GI:15795889
 KEYWORDS
 SOURCE human;
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Meyers, R.A.
 TITLE 32705; 23224; 27423; 32700; 32712; novel human g-proteins
 JOURNAL Patent: WO 0164887-A 13 07-SEP-2001;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 source
 1..1116
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 124..699
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC88255.1"
 /db_xref="GI:15795890"
 /translation="MLLGDVGKTCFLIOFKDGAFLSGTFLATVGDIFRNKVVTVDG VRVKLIWDTAGOERFSVTHAYRDAQALLLDYDITNKSSFDNIRAWLEIHEYAQR DVVIMLLGNKADMSSEVRVSEDETGLAREYGVFLETSKTKMNNVELAFALAKEL YRAGHQDEPFOIRDYVESQKRSSCCSPM"
 BASE COUNT 250 a 313 c 322 g 221 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,35e-108 Length: 1116
 Score: 1145.00 Matches: 222
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.57% Indels: 0
 DB: 6 Gaps: 0
 US-09-817-199A-2 (1-223) x AX236082 (1-1116)
 QY 2 ThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 21
 |||||
 DB 31 ACGGGCACGCCAGGCCGCTGCCACCCGGATGGGAGGCCGCCGAGCGCTCCCGGCC 90
 |||||
 QY 22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41
 |||||
 DB 91 TGCAGTCCGAGCTACGACCTCAGGCAAGGTGATGCTTCTGGGAGACACAGCGCTGCG 150
 |||||
 QY 42 LysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAla 61
 |||||
 DB 151 AAAACATGTTCTGTATCCCAATTCAAAGACGGGGCTTCTGTCCGGAACCTTCATAGCC 210
 |||||
 QY 62 ThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeu 81
 |||||

```
|||||
Db 211 ACCCTGGCATAGACTTCAGGAACAAGGTGGTGACTGGATGGCGTGAGAGCTG 270
QY 82 GlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArg 101
Db 271 CAGATCGGCACACCGCTGGCGAGAACGGTCCGAAGCGTCACCCATGCTTATTACAGA 330
QY 102 AspAlaGlnAlaLeuLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIle 121
Db 331 GATGCTCAGGCTTGCTTCTGCTGTATGACATCACCAACAATCTCTTCGACACATC 390
QY 122 ArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValIleMetLeuLeu 141
Db 391 AGGCGCTGGCTCACTGAGATTCATGATGATGCCAGAGGACGTGGTGATCATCTGCTA 450
QY 142 GlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeu 161
Db 451 GGCACACAGCGGATATGACACAGCAAGAGTATCGTTCGGAAGCGGAGAGACCTTG 510
QY 162 AlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGlu 181
Db 511 GCCAGGAGTACGGTGTCTCTTCTGGACACGAGCGCCAGACTGGCATGAATGTGGAG 570
QY 182 LeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGlu 201
Db 571 TTAGCCTTCTGGCCATCGCAAGGAAGTGAATACCGGGCGGCGATCAGCGGATGAG 630
QY 202 ProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerCysCysSer 221
Db 631 CCAGCTTCAGATCCGAGACTATGATAGATCCAGAGAGGCGCTCCAGAGGCGCTGCTC 690
QY 222 PheMet 223
Db 691 TTCATG 696

RESULT 4
AF233582 690 bp mRNA linear ROD 02-MAY-2000
LOCUS Mus musculus GTPase Rab37 (Rab37) mRNA, complete cds.
DEFINITION AF233582
ACCESSION AF233582
VERSION AF233582.1 GI:7677421
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 690)
Masuda, E.S., Luo, Y., Young, C., Shen, M., Rossi, A.B., Huang, B.C.,
Yu, S., Bennett, M.K., Pavan, D.G. and Scheller, R.H.
Rab37 is a novel mast cell specific GTPase localized to secretory
granules
JOURNAL FEBS Lett. 470 (1), 61-64 (2000)
MEDLINE 20189834
PUBMED 10722846
REFERENCE Luo, Y., Huang, B.C.B., Yu, S., Shen, M. and Masuda, E.S.
AUTHORS 2 (bases 1 to 690)
Direct Submission
TITLE Submitted (11-FEB-2000) Cell Biology, Rigel, Inc., 240 East Grand
JOURNAL Avenue, South San Francisco, CA 94080, USA
FEATURES
Source 1. .690
/organism="Mus musculus"
/db_xref="taxon:10090"
1. .690
/gene="Rab37"
1. .672
/gene="Rab37"
/codon_start=1
/product="GTPase Rab37"
/protein_id="AAF67162.1"
/db_xref="GI:7677422"
/translation="WTGTPGATAGDGAPERSPPFNVDLTGKVMLLGDSGVKTC
FLIQFDGAFLSGTFTAIIVDFRNKVVTVDGARVKLQIWDTAGOERFSVTHAYRD
```

```

AAQALLLLYDITNQSSFDNIRAWLTIETHEYAQRDVWIMLLGNKADYSSESRVISEDGET
LAREYGVPPMETSAKTMNVELAFIAIAKELKVRAGQDPDFSQIRDVVEQSKFRSS
CCSFV"
BASE COUNT 160 a 181 c 201 g 148 t
ORIGIN
Alignment Scores:
Pred. No.: 6.63e-102 Length: 690
Score: 1081.00 Matches: 209
Percent Similarity: 96.41% Conservative: 6
Best Local Similarity: 93.72% Mismatches: 8
Query Match: 94.00% Indels: 0
DB: 10 Gaps: 0
US-09-817-199A-2 (1-223) x AF233582 (1-690)
QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
Db 1 ATGACTGGCACACCGAGAGCTGTACCGCTGGGGATGGCGAGGCCCTGAGCGCTCCCG 60
QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
Db 61 CCCTTCAGCCCGGACATACGATCTACCGGCAAGGTATGCTCTTGGAGACTCGGGCGT 120
QY 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
Db 121 GGCACAAACCTGTTCTCTGATCAATTCAGAGACGGGCGCTTCCTGTCGGAACCTTCATA 180
QY 61 AlaThrValGlyIleAspPheArgAsnLysValIleThrValAspGlyValArgValLys 80
Db 181 GCACCGCTGGCATAGACTTCAGGAATAAAGTGGTGACAGTGGATGGTCAGGGGTGAAG 240
QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
Db 241 CTCAGATCTGGGACACTGCAGGACAGGAGCGCTTCGCGAGTGTGACCCATGCTTATTC 300
QY 101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
Db 301 CGAGATGCTCAGGCTTTGCTCTGTTGTATGACATCACCAACCAAGCTCTCTTTTGAAC 360
QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValIleMetLeu 140
Db 361 ATCAGGCGCTGGCTCACAGATTCATGATGATGCCAGAGGACGTGGTGATATGCTT 420
QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
Db 421 CTAGGCAACAAGCCGATGTAACACAGGAAGGTGATCCGTTCTGAGATGGAGAGACA 480
QY 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
Db 481 CTGCCAGGGAATATGCTGTTCTCTTCATGGAGACAGTCCCAAGACTGGCATGAACGTG 540
QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
Db 541 GAGTTGGCCTTTCTGGCAATTGCCAAGGAACCTGAAATACCGTCAGGGAGGACGCTGAT 600
QY 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerCysCys 220
Db 601 GAGCCAGCTTCAGATCCGAGACTATGTTGGAGTCCCAAGAGGCGCTCCAGCTGCTGC 660
QY 221 SerPheMet 223
Db 661 TCCTTTGTG 669
RESULT 5
AX236084
LOCUS AX236084
DEFINITION Sequence 15 from Patent WO0164887.
ACCESSION AX236084
VERSION AX236084.1 GI:15795891
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 576)
AUTHORS Meyers R.A.
TITLE 32705, 23224, 27423, 32700, 32712, novel human g-proteins
JOURNAL Patent: WO 0164887-A 15 07-SEP-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
Source Location/Qualifiers
1..576
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 141 a 148 c 167 g 120 t
ORIGIN
Alignment Scores:
Pred. No.: 2,48e-91 Length: 576
Score: 977.00 Matches: 191
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 84.96% Indels: 0
DB: 6 Gaps: 0
US-09-817-199a-2 (1-223) x AX236084 (1-576)
QY 33 MetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuLeuGlnPheLysAspGly 52
Db 1 ATGCTTCTGGGAGACACAGCGTCGGCAAAACATGTTCTGATCCAAATTCAGAGCGG 60
QY 53 AlaPheLeuSerGlyThrPheLeuAlaThrValGlyLeuAspPheArgAsnLysValVal 72
Db 61 GCCTTCTGTCGGAGACCTTATAGCCACCGTCGGCATAGCTTCAGGAAACAGGTGGTG 120
QY 73 ThrValAspGlyValArgValLysLeuGlnLeuTirPAspThrAlaGlyGlnGluArgPhe 92
Db 121 ACTGTGATGGCGTGAGATGAGCTGCAGATCTGGGACACCGCTGGGAGGACCGTTC 180
QY 93 ArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAspIle 112
Db 181 CGAAGCGTCACCCATGCTTATTACAGAGATGCTTCAGGCGCTTCTGCTGTATGATC 240
QY 113 ThrAsnLysSerSerPheAsnIleArgAlaTirPLeuThrGluLeuHisGluTyrAla 132
Db 241 ACCAACAATCTCTTTCGACAACATCAGGCGCTGCTCAGATTCATGATGATGCC 300
QY 133 GlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgVal 152
Db 301 CAGAGGAGCGTGTGATGCTGCTAGGCAACAGCGGATGATGACGCGAAGAGTG 360
QY 153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThr 172
Db 361 ATCCGTTCGAGACGAGAGACCTTGGCCAGGAGTACGGTGTCCCTTCCTGGAGACC 420
QY 173 SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys 192
Db 421 AGCGCAAGACTGGCATGATGAGATGGAGTATGCTTCTGCGCATCGCAAGGAACTGAA 480
QY 193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212
Db 481 TACCGGCGCGGATCAGCGGATGAGCCAGCTTCAGATCCGACATGATGTAGAGTCC 540
QY 213 GlnLysLysArgSerSerCysSerPheMet 223
Db 541 CAGAGAAGCGCTCCAGCTGCTGCTCTTCATG 573
RESULT 6
BC007681
LOCUS
DEFINITION Homo sapiens, RAB26, member RAS oncogene family, clone MGC:3503
IMAGE:3627067, mRNA, complete cds.
ACCESSION BC007681
VERSION BC007681.1 GI:14043378
KEYWORDS MGC.
SOURCE Homo sapiens.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1513)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nih.gov
Shevchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R., Lim,M., Maduro,Q.L., Mastello,C., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Snyder,B., Stantripop,S., Thomas,P.J., Tongson,E.E., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5931611.
FEATURES
Location/Qualifiers
1..1513
/organism="Homo sapiens"
/db_xref="LocusID:25837"
/db_xref="taxon:9606"
/clone="MGC:3503 IMAGE:3627067"
/tissue_type="uterus, endometrium adenocarcinoma"
/clone_lib="NIH_MGC_44"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
187..759
/codon_start=1
/product="RAB26, member RAS oncogene family"
/protein_id="AAH07681.1"
/db_xref="GI:14043379"
/translation="MLVGDGSGVGTCLLVRFKDGAFAGTGFISTVIGIDFRNKVLVDVG
VKVKLQMDTAGOERFSRVTHAYRDAHALLLYDVTNKASFDNICAWLTHEYAOH
DVALMLGNKVDSEHVRVVKREDGKLAKEYGLPFMETSAKTGLNVDLAFATLAKELK
QSRMKAPSEPRFLHDYVAREGRGASCCRP"
BASE COUNT 326 a 452 c 449 g 286 t
ORIGIN
Alignment Scores:
Pred. No.: 1,28e-70 Length: 1513
Score: 780.50 Matches: 153
Percent Similarity: 79.57% Conservatives: 30
Best Local Similarity: 66.52% Mismatches: 32
Query Match: 67.87% Indels: 16
DB: 9 Gaps: 3
US-09-817-199a-2 (1-223) x BC007681 (1-1513)
QY 6 GlyAlaValAlaThrArgAspGlyGluAla-----ProGluArgSerPro----- 20
Db 62 GGGGCGCGGCGGCGCTCCGGGACTGCGCTTTCGGGCGCGCGCGCGCGCGG 121
QY 21 ProCysSerProSer-----TyrAspLeuThrGly 30
Db 122 CCTTGAGCGCGGC-CGGCCCTCGCTTGGCGGGGTGTCGACTTCTACGACGTCGCGCTTC 180

```
QY 31 LysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuLeuGlnPheLys 50
Db 181 AAGGTCATGCTGGTGGGGAGCTGGGTGTGGGAAGACCTGCTGCTGGTGGCATTCAG 240
QY 51 AspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLys 70
Db 241 GATGGTCTTCTCTGGGGAGACCTTCATCTCCACCGTAGGCATTGACTTCCCGGAACAA 300
QY 71 ValValThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGlu 90
Db 301 GTTCTGCAGCTGGATGTTGAGGTGAAGCTGCAGATGTGGACACAGCTGGTCAGGAG 360
QY 91 ArgPheArgSerValThrHisAlaTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyr 110
Db 361 CGGTTCGCGAGTGTACCCATGCTACTACCGGAGTCTCATGCTGCTGCTGCTCTAC 420
QY 111 AspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGlu 130
Db 421 GATGTACCAACAAGGCTCTCTTTGACAACATCCAGGCTGGCTGACGAGATCCACGAG 480
QY 131 TyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGlu 150
Db 481 TACCCCGACGACGAGTGGGCTCATGCTGCTGGGAACAGGTGGACTGTGCCATGAG 540
QY 151 ArgValIleArgSerGluAspGlyGluThrLeuAlaAlaArgGluTyrGlyValPropheLeu 170
Db 541 CGTGTGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY 171 GluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGlu 190
Db 601 GAGACGAGCCCAAGAGGCGCTCAAGCTGGACTTGGCCCTTCACAGGCATAGCAAGAGG 660
QY 191 LeuLysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrVal 210
Db 661 TTGAGCAGCGCTCCATGAGGCTCCAGCGAGCGGCGCTTCCGGCTGCATGATTACGTT 720
QY 211 GluSerGlnLysLysArgSerSerCysCys 220
Db 721 AAGAGGAGGAGGTCGAGGCGGCTCTGCTGTC 750

RESULT 7
AB027137 1320 bp mRNA linear PRI 29-SEP-1999
LOCUS Homo sapiens v46133 mRNA for RAB-26, complete cds.
DEFINITION AB027137
ACCESSION AB027137
VERSION AB027137.1 GI:5931611
KEYWORDS RAB-26; v46133.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1320)
AUTHORS Miyajima,N., Seki,N., Hattori,A., Hayashi,A., Kozuma,S., Muramatsu,M. and Saito,T.
TITLE Human RAS-related protein RAB-26
JOURNAL Published Only in DataBase (1999)
REFERENCE 2 (bases 1 to 1320)
AUTHORS Miyajima,N., Seki,N., Hattori,A., Hayashi,A., Kozuma,S., Muramatsu,M. and Saito,T.
TITLE Direct Submission
JOURNAL Radiological Sciences, Genome Research Group; Inage-ku Anagawa 4-9-1, Chiba, Chiba 263-8555, Japan (E-mail:t_saito@nirs.go.jp, Tel:81-43-201-3135, Fax:81-43-251-9818)
FEATURES
source Location/Qualifiers
1..1320
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1320
/genes="v46133"
4..576
/genes="v46133"

/note="RAS-related protein"
/codon_start=1
/product="RAB-26"
/protein_id="BAA84707.1"
/db_xref="GI:5931612"
/translation="MLVGSQGVKCTLLVRFKDGAFAGTISTVGVIDPRNKVLDVVG
VKVKIQMDITAGQERSVTHAYIRDAHALLLYDVTNKASDNIQAWLITEHYAOK
DVALMLGNKVDSEHSAHRYVYKREKLEKEYGLPFMETSATKGLNLDLAATAKELK
ORSMKAPSEPRFLHDYVKREGRGASCCRP"
BASE COUNT 273 a 379 c 395 g 273 t
ORIGIN
Alignment Scores: 2,78e-68 Length: 1320
Pred. No.: 757.00 Matches: 139
Score: 757.00
Percent Similarity: 88.36% Conservative: 28
Best Local Similarity: 73.54% Mismatches: 22
Query Match: 65.83% Indels: 0
DB: 9 Gaps: 0
US-09-817-199A-2 (1-223) x AB027137 (1-1320)
QY 32 ValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuLeuGlnPheLysAsp 51
Db 1 GTCATGCTGGTGGGGAGCTCGGTGTGGGAAGACCTGTCTGTGGTGGCATTCAGGAT 60
QY 52 GlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysVal 71
Db 61 GGTGCTTCTCTGGGGGAGCTTCATCTCCACCGTAGGCATTGACTTCCCGGAACAAATT 120
QY 72 ValThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArg 91
Db 121 CTGACGCTGATGTTGAAGTGAAGTGCAGATGTGGACACAGCTGTCAGGAGCGG 180
QY 92 PheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAsp 111
Db 181 TTCGCGAGTGTACCCATGCTACTACCGGATGCTCATGCTGTGCTGCTCTACGAT 240
QY 112 IleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyr 131
Db 241 GTCAACCAAGGCTCTTTTGACACATCCAGGCTGGCTGACCGAGATCCACGAGTAC 300
QY 132 AlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArg 151
Db 301 GCCCAGCAGCAGCTGGCGCTCATGCTGCTGGGAACAAGGTGGACTCTGCCCATGAGCGT 360
QY 152 ValIleArgSerGluAspGlyGluThrLeuAlaAlaArgGluTyrGlyValPropheLeuGlu 171
Db 361 GTGGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 172 ThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeu 191
Db 421 ACCAGCGCCCAAGAGCGGCTCAAGCTGGAGTGTGGCTTTCACAGCCATAGCAAGGAGTTG 480
QY 192 LysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGlu 211
Db 481 AAGCAGCGCTCCATGAAGGCTCCACGAGCGCGCGCTTCCGGCTGCATGATTAGCTTAAG 540
QY 212 SerGlnLysLysArgSerSerCysCys 220
Db 541 AGGAGGAGGTCGAGGCGGCTCTCTGCTGC 567

RESULT 8
RNUI8771 1098 bp mRNA linear ROD 18-JUL-1995
LOCUS Rattus norvegicus Rab26 mRNA, complete cds.
DEFINITION RNUI8771
ACCESSION U18771
VERSION U18771.1 GI:619733
KEYWORDS rat.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
```

```

REFERENCE
AUTHORS 1 (bases 1 to 1098)
TITLE Wagner,A.C., Strowski,M.Z., Goke,B. and Williams,J.A.
JOURNAL Molecular cloning of a new member of the Rab protein family, Rab
MEDLINE 26, from rat pancreas
PUBMED Biochem. Biophys. Res. Commun. 207 (3), 950-956 (1995)
7864900

REFERENCE 2 (bases 1 to 1098)
AUTHORS Williams,J.A.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1994) John A. Williams, Physiology, University of
Michigan, 7744 Med. Sci. II, Ann Arbor, MI 48109, USA

FEATURES
source
1..1098
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/sex="female"
/tissue_type="pancreas"
/dev_stage="adult"
1..29
30..602
/codon_start=1
/product="Rab26"
/protein_id="AAAG69955.1"
/db_xref="GI:619734"
/translation="MLVGDGSGVGTCLLVRFKDGAGLAGTFTSTVGIDFRNKVLVDVG
MKVKLQWDTAGQERFSVTHAYYRDAHALLLYDITNKDSFDNIQAWLTIHEYAQH
DVALMLGNKVDSQERVKREDGEKLAKEYGLPFMETSAKSLGNVLAFTAIKELK
QRSTKAPSEPRFLHDYVKREGVSCRL"
603..1098
3'UTR
polyA_signal 1063..1068
BASE COUNT 267 a 277 c 306 g 248 t
ORIGIN

Alignment Scores:
Pred. No.: 4.44e-68 Length: 1098
Score: 754.00 Matches: 140
Percent Similarity: 87.89% Conservative: 27
Best Local Similarity: 73.68% Mismatches: 23
Query Match: 65.57% Indels: 0
DB: 10 Gaps: 0

US-09-817-199A-2 (1-223) x RNUI8771 (1-1098)

Qy 31 LysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLys 50
Db 24 AAGGTCATGCTGGTGGGGGATCCGGTGTGGGAAGACGTGCTGCTCGCTTCAAG 83
Qy 51 AspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLys 70
Db 84 GATGGGGCTTCCTGGCTGGTACCTTCATCTCCACTGTGGGCATCGACTTCGGGAATAA 143
Qy 71 ValValThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGlu 90
Db 144 GTTCTGGATGGTGGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 203
Qy 91 ArgPheArgSerValThrHisAlaThrTyrArgAspAlaGlnAlaLeuLeuLeuTyr 110
Db 204 CGTTCGCAAGTGTCCCATGCTGCTACTACCGGGATGTCACGCACTGTGCTGCTAC 263
Qy 111 AspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGlu 130
Db 264 GACATCCACCAAGATTCCTTCGACAAATCCAGGCTGTTGACAGAATCCAGAA 323
Qy 131 TyrAlaGlnArgAspValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGlu 150
Db 324 TATGCCACAGCAGGCTGCTGCTATGCTGTGGGAACAAGTTGACTTACTCAAGAA 383
Qy 151 ArgValIleArgSerGluAspGlyCluThrLeuAlaArgGluTyrGlyValPropheLeu 170
Db 384 CGAGTGGTAAAGAGAGATGGGGAGAAATTAAGCAAGAGATGATGGCTGCCATTTCATG 443

```

```

Qy 171 GluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGlu 190
Db 444 GAGACACGCGCAACAGCGCTCAATGTGGACTTGGCTTTTACAGCCATACCAAGGAG 503
Qy 191 LeuLysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrVal 210
Db 504 CTGAAACAAGAATCCACCAAGGCTCCAGTGAAGCGCGCTTCAGGCTGCATGACTATGTG 563
Qy 211 GluSerGlnLysLysArgSerSerCysCys 220
Db 564 AAGAGGGAGGGCCGAGGGGTCTCTGCTGT 593
RESULT 9
AF498952
LOCUS Homo sapiens small GTP binding protein RAB26 (RAB26) mRNA, complete
DEFINITION AF498952.1 GI:20379079
VERSION AF498952
KEYWORDS CDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 573)
AUTHORS Puhl,H.L. III, Ikeda,S.R. and Aronstam,R.S.
TITLE Homo sapiens RAB family small GTP binding protein RAB26
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 573)
AUTHORS Puhl,H.L. III, Ikeda,S.R. and Aronstam,R.S.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2002) cDNA Resource Center, Guthrie Research
Institute, One Guthrie Square, Sayre, PA 18840, USA

FEATURES
source
1..573
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="brain"
1..573
/gene="RAB26"
1..573
/gene="RAB26"
/feature="RAB family member"
/codon_start=1
/product="small GTP binding protein RAB26"
/protein_id="AA211100.1"
/db_xref="GI:20379080"
/translation="MLVGDGSGVGTCLLVRFKDGAGLAGTFTSTVGIDFRNKVLVDVG
VKVKLQWMDTAGQERFSVTHAYYRDAHALLLYDITNKDSFDNIQAWLTIHEYAQH
DVALMLGNKVDSQERVKREDGEKLAKEYGLPFMETSAKSLGNVLAFTAIKELK
QRSMKAPSEPRFLHDYVKREGVSCSCP"
BASE COUNT 117 a 151 c 191 g 114 t
ORIGIN

Alignment Scores:
Pred. No.: 2.41e-68 Length: 573
Score: 753.00 Matches: 138
Percent Similarity: 88.30% Conservative: 28
Best Local Similarity: 73.40% Mismatches: 22
Query Match: 65.48% Indels: 0
DB: 9 Gaps: 0

US-09-817-199A-2 (1-223) x AF498952 (1-573)

Qy 33 MetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGly 52
Db 1 ATGCTGGTGGGGAGCTCGGCTGGGAAGACCTGCTGCTGGTGGATTCAGGATGCT 60
Qy 53 AlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValVal 72
Db 61 GCTTTCCTGGGGGACCTTCATCTCCACCGTAGCATTTGACTTCCCGAACAAGTTCGTG 120

```



```

QY 73 ThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPhe 92
Db 121 GACGTGATGCTGTGAAGCTGAAGCTGCATGTTGGACACAGCTGGTCAAGGACGGTTC 180
QY 93 ArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAspIle 112
Db 181 CGCAGTGTACCATGCTTACTACCGGGATGCTCATGCTCTGCTGTGCTTACGATGTC 240
QY 113 ThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAla 132
Db 241 ACCAACAGGCCCTTTGACACATCCAGGCCCTGGCTGACCCAGATCCACAGTAGTCGC 300
QY 133 GlnArgSpValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgVal 152
Db 301 CAGCACACGTGGCTCATGCTGCTGGGAACAAGTGGACTCTGCCATGAGCGGTG 360
QY 153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThr 172
Db 361 GTGAAGAGGAGGACGGGAGAGCTGGCCACAGAGTATGGACTGCCCTTCATGGAGACC 420
QY 173 SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys 192
Db 421 AGGCCAAGACGGCCCTCAACGTGGACTTGGCCCTTCACAGCCATAGCAAAAGGAGTTGAAG 480
QY 193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212
Db 481 CAGCGTCCATGAAGCTCCACGAGCGCGCTTCGGCTGCATGATTACGTTAAGAGG 540
QY 213 GlnLysLysArgSerSerCysCys 220
Db 541 GAGGTCGAGGGCCCTCTGCTGC 564

RESULT 10
AR062279
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
BASE COUNT
ORIGIN

AR062279
Sequence 2 from patent US 5843717.
AR062279
AR062279.1 GI:5989970
Unknown.
Unclassified.
Hillman,J.L. and Guebler,K.J.
Rab protein
Patent: US 5843717-A 2 01-DEC-1998;
Location/Qualifiers
1..1340
/organism="unknown"
276 a 387 c 403 g 274 t

Alignment Scores:
Pred. No.: 1,32e-65 Length: 1340
Score: 731.00 Matches: 141
Percent Similarity: 86.73% Conservative: 29
Best Local Similarity: 71.94% Mismatches: 25
Query Match: 63.57% Indels: 2
DB: 6 Gaps: 0

US-09-817-199a-2 (1-223) x AR062279 (1-1340)

QY 26 TyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPhe 45
Db 22 TAGCAGTCCGCTTCAAGCTCATGCTGTGGGGGACTCGGGTGTGGGAAGACCTGCTG 81
QY 46 Leu-IleGlnPheLysAspGlyValAlaPheLeuSerGluThrPheIleAlaThrValClytl 65
Db 82 CTGGGTGGCATTCAGCATGGTGTCTTCTGGCGGGAGCCTTCATCTCCACCGTAGC-AT 140
QY 65 eAspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGlnIleTrpAs 85
Db 141 TGACTTCCGGACAAAGTCTTGACGTGGATGTTGTGAAGGTGAAGCTGCACATGTGGGA 200

```

```

QY 85 pThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAl 105
Db 201 CACAGCTGGTCAGGACGGTTCGCGAGTGTACCCATGCTACTACCGGAGTCTCATGC 260
QY 105 aLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLe 125
Db 261 TCTGCTGCTGCTTACGATGTACCAACAAGGCTCTCTTTGACAACATCCAGGCCCTGGCT 320
QY 125 uThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLysAl 145
Db 321 GACCGAGATCCAGGATACGCCACGACGCTGGCGCTCATGCTGCTGGGGAACAAGGT 380
QY 145 aAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTy 165
Db 381 GGACTTCGCCCATGACCGTGTGGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGTA 440
QY 165 rGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLe 185
Db 441 TGGACTGCGCTTTCATGGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 500
QY 185 uAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGluProSerPheGl 205
Db 501 AGCCATAGCAAGAGGATTGAAGCAGCGCTCCATGAAGGCTCCAGGAGGCGGCTCCG 560
QY 205 nIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
Db 561 GCTGCATGATTACGTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 606

RESULT 11
AK054846
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
BASE COUNT
ORIGIN

AK054846
3105 bp mRNA linear PRI 01-AUG-2002
Homo sapiens cDNA FLJ30284 fis, clone BRACE2002812, moderately
similar to Mus musculus GTPase Rab37 (Rab37) mRNA.
AK054846.1 GI:16549463
oligo capping; fis (full insert sequence).
Homo sapiens cerebellum cDNA to mRNA, clone_lib:BRACE2
clone:BRACE2002812.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Nishi,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
Yamamoto,T., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 3105)
Isogai,T., Otsuki,T. and Sugiyama,T.
Direct Submission
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: RAB and
HRI.
Location/Qualifiers
1..3105
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="BRACE2002812"
/tissue_type="cerebellum"
/clone_lib="BRACE2"

```



```
/note="cloning vector: pME18SFL.3"
```

BASE COUNT	680 a	905 c	898 g	622 t
ORIGIN				

Alignment Scores:

Pred. No.:	5e-63	Length:	3105
Score:	710.50	Matches:	182
Percent Similarity:	39.47%	Conservative:	13
Best Local Similarity:	36.84%	Mismatches:	26
Query Match:	61.78%	Indels:	274
DB:	9	Gaps:	5

US-09-817-199A-2 (1-223) x AK054846 (1-3105)

QY	3	GlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro-----	21
Db	375	GGAGAGCGGAGCGGAGCTCCAAGCCTGGCATGGACCTCAGAGACCCGATTCCTAC	434
QY	22	-----CysSerProSerTyrAsp-----LeuThrGlyLysValMetLeuLeuGly	36
Db	435	CAGGGAGGAGCTGGCCCTGCATTCAACGACCAGCTCTGCATAAAGACCATCTGTGGTG	494
QY	37	AspThrGlyValGlyIythrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSer	56
Db	495	GACAGTGGTGTGGGAAGAAGCGTCTCTGCTGGTTTCAGTTCGATCAGGCCAAGTTCATCC	554
QY	57	GlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValValThrValAspGly	76
Db	555	GGTCTCTTCGCCCACTGTGGCATCGGATTCACGACAAGGTGTGACTGTGGATGGC	614
QY	77	ValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThr	96
Db	615	GTGAGAGTGAAGTCAGATCTGGGACACCGCTGGGACAGAAAGCGTTCGGAAGCGTCAC	674
QY	97	HisAlaTyrrArgAspAlaGlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	116
Db	675	CATGCTATTATACAGAGATGCTCAGGCCTTGCTTCGTGTATGACATCACCAACAAATCT	734
QY	117	SerPheAsnIleArgAlaTrpLeuThrGluIleHisGluTyrrAlaGlnArgAspVal	136
Db	735	TCTTTCGAACAACTACAGGCCTGGCTCACTGAGATTTCATGATGATGCCAGGAGCAGT	794
QY	137	ValIleMetLeuLeuGlyAsnLys-----	144
Db	795	GTGATCATGCTGTAGGCAACAAAGTGAGTGGCTCGGGGCGAGGFTCAGCCAGCCCTCG	854
QY	144	-----	144
Db	855	ACTTCTCAGCCCTAGCCGCCCCATAACCACCCCAAGACAGTTATCTAGGATCCTTCC	914
QY	144	-----	144
Db	915	TGAAAAGGACTCTGTCAGCCCTCCAGCTCAGGGGTTCAGACATATCTGGAGCTTCTGCCC	974
QY	144	-----	144
Db	975	CCCATCTGCCCTTCAGGGAAGTCCAAGTTGTTCCCTTCAGAAATCAAGGGTGCCTAG	1034
QY	144	-----	144
Db	1035	TTCTCAGCCCCCATTAGACAGAGTGAACAGAGTCCAGGTCAGGGGCTTAAGAGTCAAA	1094
QY	144	-----	144
Db	1095	GGTTTAGCCCCAACTGCTGTCTATTTCGAAGACCCCTTTACCAAAGGTGAGATCCAGAG	1154
QY	144	-----	144
Db	1155	TGGGAGCTACACTGGCGCAAAACCCTGGCCCCCAGGCCAATCACACCTGCCTGCAGTCC	1214
QY	144	-----	144
Db	1215	TGGGCCACGACGAGGGCAGGCAAGCCCTGCTTCTTGGGGCAAAATATATGGGCCCGCTGG	1274

Qy	145	-----AlaSpMetSerSergl	150
Db	1275	GCGGAGCCCTCCTCCCTCCCGAGGTACGCCATTTGGGCTTGACAGCGGGATATGACGACGGA	1334
Qy	150	uAtrqValIleArgSerGLuaSpGlyCluThrLeuAlaArg	163
Db	1335	AAGAGTATCCGTTCGAAAGACGAGAGACCATTGGCCAGGGTAAGTAGTTGTCGTGGGA	1394
Qy	164	-----Gl 164	
Db	1395	CAGGGTCAAGGGTGGGGCAACCGACGCTGGCCCTGAGGACACTCTCTCCCGGSCAGGA	1454
Qy	164	uTyrgLyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaLaph	184
Db	1455	GTACGGTGTTCCTCTCTGAGACCGAGAGACCTTGCCCATGAATGTGGATTTAGCCCT	1514
Qy	184	eLeuAlaIleAlaLys-----	189
Db	1515	TCTGGGCATGCCCAA--GTGAGACTGGGCGAGGAAGGNAGTGTCCGGGGCAGGGCGCCA	1573
Qy	189	-----	189
Db	1574	CACTCCAGGAATCCACTAGTAGGGCCGCGCCCTGGCCAGCCCCTGGACACACTGCTGCTTCT	1633
Qy	189	-----	189
Db	1634	GCAGCTGAGGTCCATTGCTCTGGGAGACTGGGCCACTGGGAGAGGGGAGGGGGCGGC	1693
Qy	190	-----Gl 190	
Db	1694	TCAGCTCCTCACCCACGCCAGCCAGCCAGCCAGCCAGCCAGCCATTTGCTCTTCTTCAAGGGA	1753
Qy	190	uLeuLysTyrrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrrVa	210
Db	1754	ACTGAATAACCGGGCCGGGCATCAGCGCGATGAGCCAGCTTCCAGATCCGAGACTATGT	1813
Qy	210	lGluSerGlnLysLysArgSerSerCysSerPheMet 223	
Db	1814	AGATGCCAGAAGAAGCGTCCAGCTGCTGCTCTTCATG 1853	
RESULT	12		
LOCUS	AY061826	2536 bp mRNA linear INV 10-NOV-	
DEFINITION	Drosophila melanogaster GH21984 full length cDNA.		
ACCESSION	AY061826		
VERSION	AY061826.1 GI:16902019		
KEYWORDS	FLI CDNA.		
SOURCE	Drosophila melanogaster.		
ORGANISM	Drosophila melanogaster		
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
AUTHORS	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
	Ephydroidea; Drosophilidae; Drosophila.		
	1 (bases 1 to 2536)		
	Chapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,		
	Champè,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George,R.,		
	Gonzalez,M., Guarin,H., Li,P., Liao,G., Miranda,A., Mungall,C.J,		
	Nunco,J., Pacleb,J., Paragas,V., Park,S., Phouanavong,S., Wan		
	Yu,C., Lewis,S.E., Rubin,G.M. and Celniker,S.		
	Direct Submission		
TITLE	Submitted (02-NOV-2001) Berkeley Drosophila Genome Project,		
JOURNAL	Lawrence Berkeley National Laboratory, One Cyclotron Road,		
	Berkeley, CA 94720, USA		
COMMENT	Sequence submitted by:		
	Berkeley Drosophila Genome Project		
	Lawrence Berkeley National Laboratory		
	Berkeley, CA 94720		
	This clone was sequenced as part of a high-throughput process t		
	sequence clones from Drosophila Gene Collection 1 (Rubin et al.		
	Science 2000). The sequence has been subjected to integrity che		
	for sequence accuracy, presence of a polyA tail and contiguity		
	within 100 kb in the genome. Thus we believe the sequence to		
	reflect accurately this particular cDNA clone. However, there		

artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our web site (<http://fruitfly.berkeley.edu>) or send email to cdna@fruitfly.berkeley.edu.

```
FEATURES
  source
    1..2536
      /organism="Drosophila melanogaster"
      /strain="y; cn bw sp"
      /db_xref="taxon:7227"
      /map="78C9-78D1"
  gene
    1..2536
      /gene="CG7605"
      /note="alignment with genomic scaffold AE003594"
      /db_xref="FLYBASE:FBgn0037072"
      1172..2338
        /gene="CG7605"
        /note="Longest ORF"
        /codon_start=1
        /product="GHZ1984p"
        /protein_id="AAL27637.1"
        /db_xref="GI:16902020"
        /db_xref="FLYBASE:FBgn0037072"
        /translation="MASTAVGLGGGDPGAGPPAGSPAGSDVDVDAETT
        QARIELRRPFGDSYNPAPAPASVSASITTTTQQOQHNPSSHQSHQPSHH
        HHHHSQJLSLGSHHYDDAIAPVQSRATGTPGRPSREAMQMTAYGTDIDDDYH
        DGRWSRYDEVDMHPAPNAHQPDFTVHKHTILGDSGVGTFLVKYNTGEFRLG
        SFATGVIALTNNKVVVDGTRVKLQIWDTAGGERFSVTHAYIRDAHALLLYDTNKG
        TTYDNTRAWLGEIREYAQDVLVLGNKADCGSERQVKGREDGERLGRHNVPFMT
        SAKTGNLVELSTFAVAROLKSRGYEHDGDRGVNVDNTKARSVCAQCRNM"
  BASE COUNT      831 a      621 c      624 g      460 t
  ORIGIN
    Alignment Scores:
      Pred. No.:      7,01e-57      Length:      2536
      Score:          649,50      Matches:      132
      Percent Similarity: 75,96%      Conservative: 26
      Best Local Similarity: 63,46%      Mismatches: 48
      Query Match:      56,48%      Indels:      3
      DB:              3              Gaps:      2

us-09-817-199a-2 (1-223) x AY061826 (1-2536)

Qy 16 ProGluArgSerProProCysSerProSer---TyrAspLeuThrGlyLysValMetLeu 34
    ||| :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1698 CCCGTCACAGCGCATCAGCAGCGCCCTTCGATGATACGGTTAACAC-AAGACGATCCTG 1756
Qy 35 LeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPhe 54
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1757 CTGGGGGATTCGGCGGGGAAAGACCTCTCTCTGTCATAATACATACGGCGCAGTTC 1816
Qy 55 LeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValValThrVal 74
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1817 CGACTGGGTCTCTTCGCGCACAGTGGGCATTCGCTTACGAACAAGTGTGTGTCGTC 1876
Qy 75 AspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlnGlnGluArgPheArgSer 94
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1877 GATGAGACGCGCGTCAAGTCAATCTGGGACACAGCTGGTTCAGGAGCGATCCGGAGC 1936
Qy 95 ValThrHisAlaTyTrpArgAspAlaGlnAlaLeuLeuLeuLeuTyTrpAspIleThrAsn 114
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1937 GTTACCCACGCGTATATACGAGCGCGCAGCTCTACTGCTGCTGCTAGCAGTGACCAAC 1996
Qy 115 LysSerSerPheAsnIleArgAlaTrpLeuThrGlnIleHisGluTyAlaGlnArg 134
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1997 AAGACCACATGACAAACATTCGCGCTGGTGGCGGAGATCCGGGAGTACGCGCAGGAG 2056
Qy 135 AspValValIleMetLeuLeuGlyAsnLysAlaAspMetSer---SerGluArgValIle 153
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 2057 GACGTGGTCATCGTTTATAGGCAACAGCGGAGCTGACGCGGAGCGGAGCGGAGTGG 2116
Qy 154 ArgSerGluAspGlyGluThrLeuAlaArgGluTyrglyValProPheLeuGluThrSer 173
    :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2117 AAGCGGGAGATGGGAGCGTTTGGGGGGGAGCACACAGCTGCCCTTCTGAGACCTCG 2176
Qy 174 AlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTy 193
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2177 GCACAGCGGAGCTCAATGTGAGCTGCTCTTCACAGCGGTGGCCAGCACTAAAGAGT 2236
Qy 194 ArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyTrpValGluSerGln 213
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2237 CGCGCTAGCAGCAGCGCGATGATGAGAGTTCATGTCATGTCATGTCGTCGTCACAAAT 2296
Qy 214 LysLysArgSerSerCysSer 221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2297 ACAAGCGCGCTCGGTGTCGCGCC 2320
RESULT 13
BC019990
LOCUS
DEFINITION
  BC019990 1337 bp mRNA linear ROD 07-AUG-2002
  Mus musculus, similar to mel transforming oncogene (derived from
  cell line NK14)- RAB8 homolog, clone MGC:28462 IMAGE:4161110, mRNA,
  complete cds.
ACCESSION BC019990
VERSION BC019990.1 GI:18043408
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus.
  Eukaryota; Metazoa; Chordata; Rodentia; Sclurognathi; Muridae; Mus.
  Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1337)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-2001) National Institutes of Health, Mammalian
  Gene Collection (MGC), Cancer Genomics Office, National Cancer
  Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
  USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
  Email: cgabbs@mail.nih.gov
  Tissue Procurement: Jeffrey E. Green, M.D.
  CDNA Library Preparation: Life Technologies, Inc.
  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Baylor College of Medicine Human Genome
  Sequencing Center
  Center code: BCM-HGSC
  Web site: http://www.hgsc.bcm.tmc.edu/cdna/
  Contact: ang@bcm.tmc.edu
  Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
  Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
  Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
  through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
  Series: IRAK Plate: 37 Row: c Column: 17
  This clone was selected for full length sequencing because it
  passed the following selection criteria: Similarity but not
  identity to protein.
FEATURES
  source
    1..1337
      /organism="Mus musculus"
      /db_xref="taxon:10090"
      /map="FVB/N"
      /clone="MGC:28462 IMAGE:4161110"
      /tissue-type="Liver, normal. 5 month old male mouse."
      /clone_lib="NCI CGAP_Li9"
      /lab_host="DH10B"
      /note="Vector: pCMV-SPORT6"
      1..624
        /codon_start=1
        /product="Similar to mel transforming oncogene (derived
        from cell line NK14)- RAB8 homolog"
  CDS
```

/protein_id="AAH19990.1"
/db_xref="GI:18043409"
/translation="WAKTYDYLFKLLIGDSGVGKTCVLFRESEDAFNSTFTSTIGID
FKIRTEIDGKRIKLIQINDTAGOERFRTTAYXRGAMGIMLYVDITNEKSFQDNRW
IRNIEHASADVEKMLGNKCDVNRKQVSKERKGLADYGIKEMWTSKANINVEN
AFFTIARDIKAKMDKLENSPQSSHGKVTVEOKRTSFPRCSLL"
BASE COUNT 367 a 338 c 341 g 291 t
ORIGIN

Alignment Scores:
Pred. No.: 4.19e-44 Length: 1337
Score: 521.50 Matches: 101
Percent Similarity: 69.76% Conservative: 42
Best Local Similarity: 49.27% Mismatches: 55
Query Match: 45.35% Indels: 7
DB: 10 Gaps: 3
US-09-817-199a-2 (1-223) x BC019990 (1-1337)
QY 25 SerTyAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCys 44
Db 10 ACCTACGATTACCTGTTCAAGCTGCTGCTGATCGGGGACTCGGGGTAGGGAACCTGT 69
QY 45 PheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGly 64
Db 70 GTCCTCTTCCTGCTCCGAGGACGCGCTTC---AATCCACATTCATCTCTACCATAGGA 126
QY 65 IleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGlnIleTyr 84
Db 127 ATGACTTTAAATATTAGGACCATAGAGCTCGATGGCAAGAGATTAACTGCAGATATGG 186
QY 85 AspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGln 104
Db 187 GACACGGCCGCCAGGAGCGGTTTCGAACAATCATCAGCACGCTACTACAGGGGTGCCATG 246
QY 105 AlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTyr 124
Db 247 GGTATCATGCTGCTTACGACATATACCAATGAGAGAGTCTTTGACAAATCCGGAAATGG 306
QY 125 LeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGlyAsnLys 144
Db 307 ATTCGGAACATGAAGAGATGCTCTGAGAGCTCGAGAGATGATGATGAGGGAATAG 366
QY 145 AlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGlu 164
Db 367 TGTGATGTGAATGACAAAGACAGGTGTCGAAGGACGGGAGAAAGCTGGCACTCGAC 426
QY 165 TyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPhe 184
Db 427 TATGGATCAAGTTCATGGAGACCACTGCAAGGCCAACATTAATGTGAGAAATGCAATT 486
QY 185 LeuAlaIleAlaLysGluLeuLysTyrArg-----AlaGlyHisGlnAla 199
Db 487 TTCACCTCTTCCAGGAGATATCAAGCAAAATGAGCAAAATGGAAGGGAACACCCCG 546
QY 200 AspGluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCys 219
Db 547 CAGGGGAGCAGCCATGAGTCAAGATCACAGTGGAGCAGCAGAGAGGACCACTTCTTC 606
QY 220 ---CysSerPheMet 223
Db 607 CGGTGAGTCTCTCTG 621

RESULT 14
AB024994
LOCUS
DEFINITION
AB024994
Cicer arietinum mRNA for rab-type small GTP-binding protein,
complete cds.
ACCESSION
AB024994
VERSION
AB024994.1 GI:4586579
KEYWORDS
rab-type small GTP-binding protein.
SOURCE
Cicer arietinum (strain:ILC3279) leaf cDNA to mRNA, clone:INR134.
ORGANISM
Cicer arietinum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae;
Cicer
1 (sites)
Ichinose, Y., Tiemann, K., Schwenger-Erger, C., Toyoda, K., Hein, F.,
Hanselle, T. and Barz, W.
Genes Expressed in Ascochyta rabiei-Inoculated-Chickpea Plants and
Elicited Cell Cultures as Detected by Differential
CDNA-Hybridization
Unpublished
REFERENCE
2 (bases 1 to 1074)
AUTHORS
Ichinose, Y.
TITLE
Direct Submission
JOURNAL
Submitted (16-MAR-1999) Yuki Ichinose, Okayama University, Faculty
of Agriculture, Tsushima-naka, 1-1-1, Okayama, Okayama 700-8530,
Japan (E-mail:yuki@cc.okayama-u.ac.jp, tel: +81-86-251-8308,
Fax: +81-86-251-8308)

FEATURES
Location/Qualifiers
1..1074
/organism="Cicer arietinum"
/strain="ILC3279"
/db_xref="taxon:3827"
/clone="INR134"
/tissue_type="leaf"
107..715
/codon_start=1
/product="rab-type small GTP-binding protein"
/protein_id="BA076422.1"
/db_xref="GI:4586580"
/translation="MNPEYDYLFKLLIGDSGVGKSCLLLRADDSYLDYSISTIGVD
FKIRVEQDKTKLIQINDTAGOERFRTTAYXRGAGHIIIVYDVTDQESFNWQW
LEIDRYASENVNKLKLVGNKCDLAANKVVSSTAKAFADIEIGIPMETSAKKNATVEQ
AFMAAAAEIKNRMASQPANNRPPTVQIRGQPMNQKSGCCSI"
1074
polyA_site
/note="8 A nucleotides"
BASE COUNT 310 a 223 c 207 g 334 t
ORIGIN

Alignment Scores:
Pred. No.: 6.4e-44 Length: 1074
Score: 518.50 Matches: 97
Percent Similarity: 69.90% Conservative: 47
Best Local Similarity: 47.09% Mismatches: 55
Query Match: 45.09% Indels: 7
DB: 8 Gaps: 3

US-09-817-199a-2 (1-223) x AB024994 (1-1074)
QY 20 ProProCysSerProSerTyAspLeuThrGlyLysValMetLeuLeuGlyAspThrGly 39
Db 101 CCCGTATGAATCCGCAATATGACTATTGTTCAAGCTTTGTTGATGGAGATCTCGT 160
QY 40 ValGlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPhe 59
Db 161 GTGGCAAGTCATGCTCTCTCTGAGTGTGCTGATGATTCATACCTTGAC---AGCTAT 217
QY 60 IleAlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgVal 79
Db 218 ATCAGTACAAATGGAGTGGACTTTAAATTCGCACGTGTTGAGCAAGACGGGAAGACATT 277
QY 80 LysLeuGlnIleTyrAspThrAlaGlyGlnArgPheArgSerValThrHisAlaTyr 99
Db 278 ARACTTCAATTTGGGACACTGCTGTCAGAACGTTTCGGGACTATCTACTACGACTAC 337
QY 100 TyrArgAspAlaGlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 119
Db 338 TATCGTGGGCTCATGGCATAATTTGTTTATGATGTCACCTACCAAGAGAGACTTAAC 397
QY 120 AsnIleArgAlaTyrLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMet 139
Db 398 AATGTTAAGCAGTGGTGAATGAATGACCGTTATGCAAGTGAATAATGTAACAAGCTT 457

```

QY 140 LeuLeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGlu 159
||||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db 458 CTAGTTGGAAAGAGTGTATCTCGCGCAAAATAAAGTTGTCTCTGAAACAGCAAG 517

QY 160 ThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsn 179
||| ||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db 518 GCATTTGCAGATGAATTCGAATTCCTTCATGGAACACTAGTCTAAATAATGCCACCAAC 577

QY 180 ValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGln--- 198
||||| ||||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db 578 GTGGAACAGCTTTCATGGCCATGGCTGCTGAAATATAAAACAGAGATGGCAAGTCAACCA 637

QY 199 -----AlaAspGluProSerPheGlnIleArgAspTyrValGluSerGlnLysLys 215
||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db 638 GCGAACAAATCCAGGCTCCAACAGTCCAAATTCGAGGACGAGCAATGAACCAAAAG--- 694

QY 216 ArgSerSerCysCysSer 221
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 695 ---TCGTGTTGCTGCTCG 709

RESULT 15
DDISASIA
LOCUS DDISASIA 2145 bp mRNA linear INV 27-APR-1993
DEFINITION D.discoideum GTP-binding protein (SAS1) gene, complete cds.
ACCESSION M34456
VERSION M34456.1 Gi:167874
KEYWORDS GTP-binding protein.
SOURCE D.discoideum, cDNA to mRNA.
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE
1 (bases 1 to 2145)
AUTHORS Saxe,S.A. and Kimmel,A.R.
TITLE SAS1 and SAS2, GTP-binding protein genes in Dictyostelium
discoideum with sequence similarities to essential genes in
Saccharomyces cerevisiae
Mol. Cell. Biol. 10 (5), 2367-2378 (1990)
JOURNAL 90220623
MEDLINE 2109188
PUBMED
FEATURES
source
1..2145
Location/Qualifiers
/organism="Dictyostelium discoideum"
/db_xref="taxon:44689"
1095..1706
/note="GTP-binding protein (SAS1)"
/codon_start=1
/db_xref="GI:167875"
/translation="MTSPATNKPAAYDFLVKLLLLIGDSGVKSCILLRFSQSGSTPSF
IATIGDKIRITIELEGRKIKQIWDTAGQERFRTITAYRGAMGILLVYDVTDEKS
FGSIRNWINIIEQHASDVNKMIGNKDMTEKVKYVDSRGSRLADEYGIKFLTSK
NSVNVEEAFIGLAKDKIKRMIDTPNDPDHTICITPNKKNTCC"
BASE COUNT 929 a 249 c 238 g 727 t 2 others
ORIGIN

Alignment Scores:
Pred. No.: 1.99e-43 Length: 2145
Score: 517.50 Matches: 97
Percent Similarity: 71.36% Conservative: 50
Best Local Similarity: 47.09% Mismatches: 54
Query Match: 45.00% Indels: 5
DB: 3 Gaps: 3

US-09-817-199a-2 (1-223) x DDISASIA (1-2145)
QY 15 AlaProGluArgSerProProCysSerProSerTyrAspLeuThrGlyLysValMetLeu 34
:::||||| ::|:|::| ::|:|::| ::|:|::| ::|:|::| ::|:|::|
Db 1101 TCTCCAGCACAAATAAACAGCA-----GCTACGATTTTGTAGTTAAATACATTITA 1154

QY 35 LeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPhe 54
:::||||| ::|:|::| ::|:|::| ::|:|::| ::|:|::| ::|:|::|
Db 1155 ATGTGGGATAGTGGTGTAGAAAGTCATGCTTTTATTACGTTTTCGTGATGTTCTTTC 1214

QY 55 LeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValValThrVal 74
:::||||| ::|:|::| ::|:|::| ::|:|::| ::|:|::| ::|:|::|

```

Search completed: November 17, 2002, 17:32:49
Job time : 2657 secs

```

Db 1215 ---ACACCAAGTTTCATCGCTACTATTGGTATCGATTTCAAAATTCGTAACATTGAATTA 1271

QY 75 AspGlyValArgValLysLeuGlnIleTyrPaspThrAlaGlyGlnGluArgPheArgSer 94
:::||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db 1272 GAGGGTAAAGAATTAATAATTACAATAATTCGGACACTGCAGGTCAAGAAAGATTCAGAACT 1331

QY 95 ValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsn 114
:::||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db 1332 ATCACTACAGCACTATCGCGGTATGGGTATGCTATGTTGTTTATGTTATGTTCACTGAT 1391

QY 115 LysSerSerPheAspAsnIleArgAlaTyrPLeuThrGluIleHisGluTyrAlaGlnArg 134
:::||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db 1392 GAAAAATCTTTTGGTAGCATTAGAAAATGGATTAGAAAATATCGAGCAACATGCTTCAGAC 1451

QY 135 AspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgValIleArg 154
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1452 TCAGTTAATAAATGTTAATCGGTAAATAATGTGATATGATGATGATGATGATGATGAT 1511

QY 155 SerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAla 174
||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db 1512 AGCTCAAGAGGTAATAATCACTTGCAGACGAATATGGTATTAAATTTTAGAAACTTCTGCC 1571

QY 175 LysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArg 194
||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db 1572 AAAACACAGTGAATAATGTAGAGGAAGCGCTTTATTGGTTTGTAGCAAAAGATATTAACAAACGT 1631

QY 195 AlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSerGlnLys 214
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1632 ATGATTGATACACCAATATGATCTGATCATACCATAT-----TGCATTACTCCAACAAT 1685

QY 215 LysArgSerSerCysCys 220
|||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db 1686 AAGAAAAAATACTTGTGT 1703

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2002, 14:38:33 ; Search time 25 Seconds
(without alignments)
369.969 Million cell updates/sec

Title: US-09-817-199a-2

Perfect score: 1150

Sequence: 1 MTGTPGAVATRDGEAPERSP.....FOIRDYVESQKRKSCCSFPM 223

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1150	100.0	223	1 RB37_HUMAN	Q96ax2 homo sapien
2	1081	94.0	223	1 RB37_MOUSE	Q9jkm7 mus musculu
3	753	65.5	190	1 RB26_HUMAN	Q9ulw5 homo sapien
4	745	64.8	190	1 RB26_RAT	P51156 rattus norv
5	517.5	45.0	203	1 SAS2_DICDI	P20791 dictyosteli
6	516	44.9	206	1 RAB8_MOUSE	P55258 mus musculu
7	513.5	44.7	207	1 RAB8_HUMAN	P24407 homo sapien
8	512.5	44.6	215	1 RAB1_BETVU	Q39433 beta vulgar
9	512	44.5	210	1 RAB8_DISOM	P22128 discoppyge o
10	507	44.1	258	1 ARA5_ARATH	P28188 arabidopsis
11	502.5	43.7	207	1 RB8B_RAT	P70550 rattus norv
12	501	43.6	207	1 RB8B_HUMAN	Q29230 homo sapien
13	500	43.5	202	1 RIC1_ORYSA	P40392 oryza sativ
14	498	43.3	208	1 SAS1_DICDI	P20790 dictyosteli
15	495	43.0	203	1 YPT2_MAIZE	Q05737 zea mays (m
16	495	43.0	216	1 ARA3_ARATH	P28186 arabidopsis
17	494	43.0	200	1 RB10_CANFA	P24409 canis famil
18	494	43.0	200	1 RB10_HUMAN	O88386 homo sapien
19	490.5	42.7	200	1 RAO1_DISOM	P22127 discoppyge o
20	487	42.3	200	1 YPT2_SCHPO	Q39571 chlamydomon
21	486	42.3	203	1 YPT1_CHLRE	P16976 zea mays (m
22	479.5	41.7	208	1 YPT1_MAIZE	P16976 zea mays (m
23	478	41.6	203	1 YPT1_VOLCA	P31584 volvox cart
24	477.5	41.5	205	1 RAB1_LYNST	Q05974 lymnaea sta
25	476	41.4	217	1 YPT2_VOLCA	P36861 volvox cart
26	473.5	41.2	203	1 RB13_HUMAN	P51153 homo sapien
27	472	41.0	201	1 RB1B_RAT	P10536 rattus norv
28	465.5	40.5	202	1 RAB1_DISOM	P22125 discoppyge o
29	464.5	40.4	203	1 RYLL_YARLI	P41924 yarrowia li
30	461	40.1	200	1 RB10_RAT	P35281 rattus norv
31	459.5	40.0	203	1 YPT1_NEUCR	P33723 neospora
32	458.5	39.9	205	1 RB1A_HUMAN	P11476 homo sapien
33	456.5	39.7	205	1 RB1A_RAT	P05711 rattus norv

RESULT 1

```
RB37_HUMAN ID RB37_HUMAN STANDARD; PRT; 223 AA.
AC Q96AX2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Ras-related protein Rab-37.
GN RAB37.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Strausberg R.;
RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
CC -|- SUBCELLULAR LOCATION: Secretory granules (By similarity).
CC -|- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; BC016615; AAL16615.1;
DR InterPro; IPR001806; Ras_trnsfrmg.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR TIGRfams; TIGR00231; small_Gtp; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 36 43 GTP (BY SIMILARITY).
FT NP_BIND 85 89 GTP (BY SIMILARITY).
FT NP_BIND 143 146 GTP (BY SIMILARITY).
FT DOMAIN 59 67 EFECTOR REGION (BY SIMILARITY).
FT LIPID 219 219 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 220 220 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 223 AA; 24815 MW; 5A7A4887BCB84A8 CRC64;
```

Query Match 100.0%; Score 1150; DB 1; Length 223;
Best Local Similarity 100.0%; Pred. No. 2.4e-94;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMILGDTGVGKTCFLIQFKDGAFLSGTFI 60

Db 1 MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMILGDTGVGKTCFLIQFKDGAFLSGTFI 60

Qy 61 ATVGIDFENKVVTVGVKVLQIWDTAGQERFRSVTHAYYRDAQAALLLLYDITNKSSFDN 120

Db 61 ATVGIDFENKVVTVGVKVLQIWDTAGQERFRSVTHAYYRDAQAALLLLYDITNKSSFDN 120

```
QY 121 IRAWLTIHEYAQRDVVIMLGNKADSMSSERVIRSDGETLAREYGVFPLETSAGTKGMV 180
|||||
Db 121 IRAWLTIHEYAQRDVVIMLGNKADSMSSERVIRSDGETLAREYGVFPLETSAGTKGMV 180
|||||
QY 181 ELAFLAIAKELKYRAGHQADEFSQIRDVESQKRSSCCSF 223
|||||
Db 181 ELAFLAIAKELKYRAGHQADEFSQIRDVESQKRSSCCSF 223
|||||

RESULT 2
RB37_MOUSE
ID RB37_MOUSE STANDARD; PRT; 223 AA.
AC Q9JRM7;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Ras-related protein Rab-37.
GN Rab37.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20189834; PubMed=10722846;
RA Masuda E.S., Luo Y., Young C., Shen M., Rossi A.B., Huang B.C., Yu S.,
RA Bennett M.K., Pavan D.G., Scheller R.H.;
RT "Rab37 is a novel mast cell specific GTPase localized to secretory
RT granules.";
RL FEBS Lett. 470:61-64(2000).
CC -!- SUBCELLULAR LOCATION: Secretory granules.
CC -!- TISSUE SPECIFICITY: Specifically expressed in the bone marrow mast
CC cells.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF233582; AAF67162.1; -.
DR HSSP; P01112; 1PLJ.
DR MGD; MGI:1929945; Rab37.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsmfmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 36 43 GTP (BY SIMILARITY).
FT NP_BIND 85 89 GTP (BY SIMILARITY).
FT NP_BIND 143 146 GTP (BY SIMILARITY).
FT DOMAIN 59 67 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 219 219 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 220 220 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 223 AA; 24716 MW; E72013621C12C605 CRC64;

Query Match 94.0%; Score 1081; DB 1; Length 223;
Best Local Similarity 93.7%; Pred. No. 2.9e-88;
Matches 209; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MTGTPGAVATRGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQKDGAFISGTFI 60
|||||
Db 1 MTGTPGAATAGDGEAPERSPFPSPNYDLTGKVMLLGDSGVGKTCFLIQKDGAFISGTFI 60
|||||

QY 61 ATVGIDFRNKVTVDCGVRYKLOIWDTAGOERFSVTHAYYRDAQALLLYDITNKSSFDN 120
|||||
Db 61 ATVGIDFRNKVTVDCGVRYKLOIWDTAGOERFSVTHAYYRDAQALLLYDITNKSSFDN 120
|||||
```

```
QY 121 IRAWLTIHEYAQRDVVIMLGNKADSMSSERVIRSDGETLAREYGVFPLETSAGTKGMV 180
|||||
Db 121 IRAWLTIHEYAQRDVVIMLGNKADSVSERSIRSDGETLAREYGVFPMEITSAGTKGMV 180
|||||
QY 181 ELAFLAIAKELKYRAGHQADEFSQIRDVESQKRSSCCSF 223
|||||
Db 181 ELAFLAIAKELKYRAGQDPDFSFQIRDVESQKRSSCCSFV 223
|||||

RESULT 3
RB26_HUMAN
ID RB26_HUMAN STANDARD; PRT; 190 AA.
AC Q9ULW5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ras-related protein Rab-26.
GN Rab26.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Fetal brain;
RX MEDLINE=20496221; PubMed=11043516;
RA Seki N., Yoshikawa T., Hattori A., Miyajima N., Muramatsu M.,
RA Saito T.;
RT "cDNA cloning of a human RAB26-related gene encoding a Ras-like
RT GTP-binding protein on chromosome 16p13.3 region.";
RL J. Hum. Genet. 45:309-314(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RA Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN BRAIN.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AB027137; BAA84707.1; -.
DR EMBL; BC007681; AAH07681.1; -.
DR HSSP; P36017; 1ER0.
DR Genew; HGNC:14259; RAB26.
DR MIM; 605455; -.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsmfmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation.
FT NP_BIND 4 11 GTP (BY SIMILARITY).
FT NP_BIND 53 57 GTP (BY SIMILARITY).
FT NP_BIND 111 114 GTP (BY SIMILARITY).
FT DOMAIN 27 35 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 187 187 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 188 188 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 190 AA; 21354 MW; 517387B7C0B3291C CRC64;

Query Match 65.5%; Score 753; DB 1; Length 190;
Best Local Similarity 73.4%; Pred. No. 2e-59;
Matches 138; Conservative 28; Mismatches 22; Indels 0; Gaps 0;
```

QY 33 MLGDTGVGKTCFLQFKDGAFLSGTFTATVGIIDFRNKKVTVTVGVVVKVQLQINDTAGQERF 92
 Db 1 MLVGDGSGVKTCLLVRFKDGAFLAGTFTSTVGIIDFRNKKVTVTVGVVVKVQLQINDTAGQERF 60
 QY 93 RSVTHAYYRDAQALLLLYDITNKKSFNDINRAWLTEIHEYAQRDQVVMLLGNKADMSSEV 152
 Db 61 RSVTHAYYRDAQALLLLYDITNKKSFNDINRAWLTEIHEYAQRDQVVMLLGNKADMSSEV 120
 QY 153 IRSEGGTELAREYGVFPFLETSKTMNVLAFATLAKELKYRAGHQADSPFOIRDYVES 212
 Db 121 VREDEGEKLAKEYGLPFPMETSAKSLGNVLDLAFATLAKELKQKSTKAPSEPRFLHDYVKR 180
 QY 213 QKRRSSCC 220
 Db 181 EGRGASCC 188

RESULT 4
 RB26_RAT
 ID RB26_RAT STANDARD; PRT; 190 AA.
 AC P51156;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Ras-related protein Rab-26.
 GN RAB26.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Pancreas;
 RX MEDLINE=95169156; PubMed=7864900;
 RA Wagner A.C.C., Srowski M.Z., Goeko B., Williams J.A.;
 RT "Molecular cloning of a new member of the Rab protein family, Rab 26,
 from rat pancreas";
 RL Biochem. Biophys. Res. Commun. 207:950-956(1995).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN PANCREAS, KIDNEY, BRAIN,
 CC SUBMANDIBULAR GLAND, AND LUNG.
 CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U18771; AAA69955.1; -
 CC HSSP; P05713; 3RAB.
 CC InterPro; IPR003579; GTPase_Rab.
 CC InterPro; IPR001806; Ras_trnsfrmg.
 CC InterPro; IPR005225; Small_GTP.
 CC Pfam; PF00071; ras; 1.
 CC SMART; SM00175; RAB; 1.
 CC TIGRFAMs; TIGR00231; small_GTP; 1.
 CC GTP-binding; Lipoprotein; Prenylation.
 FT NP_BIND 4 11 GTP (BY SIMILARITY).
 FT NP_BIND 53 57 GTP (BY SIMILARITY).
 FT NP_BIND 111 114 GTP (BY SIMILARITY).
 FT DOMAIN 27 35 EFFECTOR REGION (BY SIMILARITY).
 FT LIPID 187 187 GERANYL-GERANYL (BY SIMILARITY).
 FT LIPID 188 188 GERANYL-GERANYL (BY SIMILARITY).
 SQ SEQUENCE 190 AA; 21457 MW; 96B2D1661F363031 CRC64;

Query Match 64.8%; Score 745; DB 1; Length 190;
 Best Local Similarity 73.4%; Pred. No. 1e-58;
 Matches 138; Conservative 27; Mismatches 23; Indels 0; Gaps 0;

QY 33 MLGDTGVGKTCFLQFKDGAFLSGTFTATVGIIDFRNKKVTVTVGVVVKVQLQINDTAGQERF 92
 Db 1 MLVGDGSGVKTCLLVRFKDGAFLAGTFTSTVGIIDFRNKKVTVTVGVVVKVQLQINDTAGQERF 60
 QY 93 RSVTHAYYRDAQALLLLYDITNKKSFNDINRAWLTEIHEYAQRDQVVMLLGNKADMSSEV 152
 Db 61 RSVTHAYYRDAQALLLLYDITNKKSFNDINRAWLTEIHEYAQRDQVVMLLGNKADMSSEV 120
 QY 153 IRSEGGTELAREYGVFPFLETSKTMNVLAFATLAKELKYRAGHQADSPFOIRDYVES 212
 Db 121 VREDEGEKLAKEYGLPFPMETSAKSLGNVLDLAFATLAKELKQKSTKAPSEPRFLHDYVKR 180
 QY 213 QKRRSSCC 220
 Db 181 EGRGASCC 188

Db 1 MLVGDGSGVKTCLLVRFKDGAFLAGTFTSTVGIIDFRNKKVTVTVGVVVKVQLQINDTAGQERF 60
 QY 93 RSVTHAYYRDAQALLLLYDITNKKSFNDINRAWLTEIHEYAQRDQVVMLLGNKADMSSEV 152
 Db 61 RSVTHAYYRDAQALLLLYDITNKKSFNDINRAWLTEIHEYAQRDQVVMLLGNKADMSSEV 120
 QY 153 IRSEGGTELAREYGVFPFLETSKTMNVLAFATLAKELKYRAGHQADSPFOIRDYVES 212
 Db 121 VREDEGEKLAKEYGLPFPMETSAKSLGNVLDLAFATLAKELKQKSTKAPSEPRFLHDYVKR 180
 QY 213 QKRRSSCC 220
 Db 181 EGRGASCC 188

RESULT 5
 SAS2_DICDI
 ID SAS2_DICDI STANDARD; PRT; 203 AA.
 AC P20791;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE GTP-binding protein SAS2.
 GN SASB OR SAS2.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90220623; PubMed=2109188;
 RA Saxe S.A., Kimmel A.R.;
 RT "SAS1 and SAS2, GTP-binding protein genes in Dictyostelium discoideum
 RT with sequence similarities to essential genes in Saccharomyces
 RT cerevisiae";
 RL Mol. Cell. Biol. 10:2367-2378(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89209367; PubMed=3149563;
 RA Saxe S.A., Kimmel A.R.;
 RT "Genes encoding novel GTP-binding proteins in Dictyostelium";
 RL Dev. Genet. 9:259-265(1988).
 CC -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
 CC TRAFFIC (BY SIMILARITY).
 CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M34456; AAA33248.1; -
 CC PIR; B34716; B34716.
 CC PIR; B61571; B61571.
 CC HSSP; P05713; 3RAB.
 CC DictyDb; DD05021; sasB.
 CC InterPro; IPR003579; GTPase_Rab.
 CC InterPro; IPR001806; Ras_trnsfrmg.
 CC InterPro; IPR005225; Small_GTP.
 CC Pfam; PF00071; ras; 1.
 CC PRINTS; PR00449; RASTRNSFRMG.
 CC SMART; SM00175; RAB; 1.
 CC TIGRFAMs; TIGR00231; small_GTP; 1.
 KW GTP-binding; Lipoprotein; Palmitate.
 FT NP_BIND 22 29 GTP (BY SIMILARITY).
 FT NP_BIND 70 74 GTP (BY SIMILARITY).
 FT NP_BIND 128 131 GTP (BY SIMILARITY).
 FT LIPID 202 202 PALMITATE (BY SIMILARITY).
 FT LIPID 203 203 PALMITATE (BY SIMILARITY).
 SQ SEQUENCE 203 AA; 22642 MW; 33EFB2F451ECAC96 CRC64;

Query Match 45.0%; Score 517.5; DB 1; Length 203;
Best Local Similarity 47.1%; Pred. No. 1.3e-38;
Matches 97; Conservative 50; Mismatches 54; Indels 5; Gaps 3;

QY 15 APERSPPSPSYDLTGKVMILGDTGKTCFLIQKDGAFSLGTFIATVGIDFRNKVVTV 74
DB 3 SPATNKFA--AYDFLVKLLIGSGVKSCLLRFSDGSP-TPSFIATIGDKIRTIEL 59
QY 75 DGVVKLIQIWDTAGQERFSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTHEHYAQR 134
DB 60 EGKRIKLIQIWDTAGQERFRITAYYRGAMGILLVYDTEKSFGRNWRNIEHQASD 119
QY 135 DVVIMLGNKADMSRVSIRSEGETIAREYGVFFLETSAGTGMNVELAFIATAKELKYR 194
DB 120 SVNKLGNKCDMTKKVDSRSKSLADEYGIKFLTSAKNSVNVBEAFIGLAKDIKKR 179
QY 195 AGHQADEPFIQIRDYVESQKRRSSCC 220
DB 180 MIDTPNDPDHTI--CITPNKKNTCC 203

RESULT 6

RAB8_MOUSE STANDARD; PRT; 206 AA.
ID RAB8_MOUSE
AC P55258;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Ras-related protein Rab-8 (Oncogene c-mel).
GN RAB8 OR MEL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91360267; PubMed=1886711;
RA Nimmo E.R., Sanders P.G., Padua R.A., Hughes D., Williamson R.,
Johnson K.J.;
RT "The MEL gene: a new member of the RAB/YPT class of RAS-related
genes";
RL Oncogene 6:1347-1351(1991).
CC -!- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
CC NEUROTRANSMITTER RELEASE.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S53270; AAB19682.1; -
CC HSSP; P05713; 3RAB
CC MGD; MGI:96960; Mel.
CC InterPro; IPR003579; GTPase_Rab.
CC InterPro; IPR001230; Prenyl_site.
CC InterPro; IPR001806; Ras_trnsmfng.
CC InterPro; IPR005225; Small_GTP.
CC Pfam; PF00071; ras; 1.
CC PRINTS; PR00449; RASTNSFRMNG.
CC SMART; SM00175; RAB; 1.
CC TIGRfam; TIGR00231; small_GTP; 1.
KW GTP-binding; Prenylation; Lipoprotein; Protein transport;
KW Proto-oncogene.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 203 203 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 206 AA; 23557 MW; 49D832725D662942 CRC64;

Query Match 44.9%; Score 516; DB 1; Length 206;
Best Local Similarity 49.5%; Pred. No. 1.7e-38;
Matches 101; Conservative 41; Mismatches 56; Indels 6; Gaps 3;
QY 25 SYDLTKVMILGDTGKTCFLIQKDGAFSLGTFIATVGIDFRNKVVTVGVVVKLIQIW 84
DB 4 TYDLFKLLIGSGVKGTCVLFREFSDAF-NSTFISTIGDKIRTIELDKRIKLIQIW 62
QY 85 DTAGQERFSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTHEHYAQRDVVIMLGNK 144
DB 63 DTAGQERFRITAYYRGAMGIMLVYDITNEKSFNIRNWRNIEEHASADVEKMILGNK 122
QY 145 ADMSSRVSIRSEGETIAREYGVFFLETSAGTGMNVELAFIATAKELKYRAGHO---AD 200
DB 123 CDVNDKRVQSKERGEKALDYGIKFMETSAKANINVENAFETLARDIKAKMDKNWKATAA 182
QY 201 EPSFIQIRDYVESQKRRSSC-CSEFM 223
DB 183 GSSHGKIVTEQOKRTSFRCSSL 206

RESULT 7

RAB8_HUMAN STANDARD; PRT; 207 AA.
ID RAB8_HUMAN
AC P24407;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Ras-related protein Rab-8 (Oncogene c-mel).
GN RAB8 OR MEL.
OS Homo sapiens (Human), and
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606, 9615;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RX MEDLINE=94124602; PubMed=8294494;
RA Zahraoui A., Joberty G., Arpin M., Fontaine J.J., Hellio R.,
Tavitian A., Louvard D.;
RT "A small rab GTPase is distributed in cytoplasmic vesicles in non
RT polarized cells but colocalizes with the tight junction marker ZO-1
RT in polarized epithelial cells";
RL J. Cell Biol. 124:101-115(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RX MEDLINE=91360267; PubMed=1886711;
RA Nimmo E.R., Sanders P.G., Padua R.A., Hughes D., Williamson R.,
Johnson K.J.;
RT "The MEL gene: a new member of the RAB/YPT class of RAS-related
RT genes";
RL Oncogene 6:1347-1351(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=Human; TISSUE=Skin;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=C.familiaris; STRAIN=Cocker spaniel;
RX MEDLINE=91061765; PubMed=2123294;
RA Chavrier P., Vingron M., Sander C., Simons K., Zerial M.;
RT "Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell
RT line";
RL Mol. Cell. Biol. 10:6578-6585(1990).
CC -!- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
CC NEUROTRANSMITTER RELEASE.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the ENBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

```

CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M38391; AAA49232.1; -
CC PIR: B38625; B38625.
CC HSP: P05713; 3RAB.
CC InterPro: IPR003579; GTPase_Rab.
CC InterPro: IPR001230; Prenyl_site.
CC InterPro: IPR001806; Ras_trnsfrmng.
CC InterPro: IPR005225; Small_GTP.
CC Pfam: PF00071; ras; 1.
CC PRINTS: PR00449; RASTRNSFRMG.
CC SMART: SM00175; RAB; 1.
CC TIGRFAMs: TIGR00231; small_GTP; 1.
CC GTP-binding; Prenylation; Lipoprotein.
CC NP_BIND 15 22 GTP (BY SIMILARITY).
CC FT NP_BIND 63 67 GTP (BY SIMILARITY).
CC FT NP_BIND 121 124 GTP (BY SIMILARITY).
CC FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
CC FT LIPID 207 207 GERANYL-GERANYL (BY SIMILARITY).
CC FT SEQUENCE 210 AA; 24164 MW; 0F73EDB08B9EEA CRC64;

Query Match 44.5%; Score 512; DB 1; Length 210;
Best Local Similarity 48.6%; Pred. No. 4e-36;
Matches 102; Conservative 44; Mismatches 50; Indels 14; Gaps 4;

QY 25 SYDLTGKVMLLGDTGVGKTCFLIQKGAFSGTFTATVIGIDFRNKKVTVGVGVKQLQIW 84
Db 4 TYDLEKLLIGDSGVGKTCFLRFSEDAF-NTFTSTIGIDFKIRTVELDGKKIKLQIW 62

QY 85 DTAGQGRFRSVTHAYYRDAQALLLYDITNKSFDNIRAWLTFIHEYAQRDQVIMLLGNK 144
Db 63 DTAGQGRFRFTITAYYRGAMGIMKVVVDITNKSFDNIRAWLTFIHEYAQRDQVIMLLGNK 122

QY 145 ADMSSSRVSRSEGTAREYGVPELETSAKTCGMVLEAFALAIKELKVRAGHQADPEPF 204
Db 123 CDNWRKQVSKERGEKLAIDYIGKFLTSAKSSINVEEAFITLARDIMTKLNKMNENSL 182

QY 205 QIRDYVE-----SQKR--SSCCSEF 223
Db 183 Q--EAVDKLSPKPKPSOKKQLSPRCSLL 210

RESULT 10
ID ARAS_ARATH STANDARD; PRT; 258 AA.
AC P28188;
DT 01-DEC-1992 (Rel. 24, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ras-related protein ARA-5.
GN ARA-5 OR AF1G02130 OR T7123.6.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,

```

```

RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Millscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
RN [2]
RP SEQUENCE OF 64-258 FROM N.A.
RC STRAIN=Various strains; TISSUE=Leaf;
RX MEDLINE=92084144; PubMed=1748311;
RA Anal T., Hasegawa K., Watanabe Y., Uchimiya H., Ishizaki R.,
RA Matsui M.;
RT "Isolation and analysis of cDNAs encoding small GTP-binding proteins
RT of Arabidopsis thaliana.";
RL Gene 108:259-264(1991).
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U89959; AAC24370.1; -
CC EMBL: D01027; BAA00832.1; -
CC PIR: PS0279; PS0279.
CC HSP: P05713; 3RAB.
CC InterPro: IPR003579; GTPase_Rab.
CC InterPro: IPR001806; Ras_trnsfrmng.
CC InterPro: IPR005225; Small_GTP.
CC Pfam: PF00071; ras; 1.
CC PRINTS: PR00449; RASTRNSFRMG.
CC SMART: SM00175; RAB; 1.
CC TIGRFAMs: TIGR00231; small_GTP; 1.
CC GTP-binding; Lipoprotein; Prenylation; Multigene family.
CC NP_BIND 70 77 GTP (BY SIMILARITY).
CC FT NP_BIND 118 122 GTP (BY SIMILARITY).
CC FT NP_BIND 176 179 GTP (BY SIMILARITY).
CC FT DOMAIN 92 100 EFFECTOR REGION (BY SIMILARITY).
CC FT LIPID 255 255 GERANYL-GERANYL (BY SIMILARITY).
CC FT LIPID 256 256 GERANYL-GERANYL (BY SIMILARITY).
CC SEQUENCE 258 AA; 29169 MW; 34D278A0C504638F CRC64;

Query Match 44.1%; Score 507; DB 1; Length 258;
Best Local Similarity 45.5%; Pred. No. 1.4e-37;
Matches 95; Conservative 50; Mismatches 56; Indels 8; Gaps 3;

QY 18 RSPGCPSPYDLTGKVMLLGDTGVGKTCFLIQKGAFSGTFTATVIGIDFRNKKVTVGV 77
Db 52 RSKTNPEVDYLFKLLLLIGDSGVGSKLLRSDSYVE-SYSTIGVDKFTVEQDGK 110

QY 78 RYKLIQWDTAGQGRFRSVTHAYYRDAQALLLYDITNKSFDNIRAWLTFIHEYAQRDQV 137
Db 111 TIKLIQWDTAGQGRFRFTITAYYRGAMGIIIVDVTDEESFNNVQWLSEIDRYASDNV 170

QY 138 IMLGNKADMSRVSIRSEGTAREYGVPELETSAKTCGMVLEAFALAIKELKVR--- 194
Db 171 KLLVGNKSLDTENRAIPYETAFAKAFADEIGIPMETSAKDATNVEQAFMAASIKERMA 230

QY 195 --AGHQADEPFOIRDYVESQKRSKSCS 221
Db 231 QPAGNNARPPTVQIRGPVAK--NGCCS 257

```

[illegible]

Matches 95; Conservative 49; Mismatches 50; Indels 10; Gaps 3;

QY 23 SPSYDLTGKVMILGDTGCGKTCFLQKDGAFLSGTFIATVGIDFRNKVYTVDGVVRVKLQ 82
DB 9 SAAADYLIKLLLLIGDSGVKSCLLLRFSDF-TPSFITIGIDFKIRITIELEGGRIKIQ 67
QY 83 IWDTAGCQERFRSVTHAYYRDAQALLLLYDITNKKSSFDNIRAMLTETHEYAQRDVMIMLG 142
DB 68 IWDTAGCQERFRITATYYRGAGILLVYDVTDEKSFGRNIRNIRTEQHATDSVNMKMLIG 127
QY 143 NKADMSERVIRSEDEGTALAREYGVFPLETSAGTKGNVELAFLAIAKELKYRAGHQADSP 202
DB 128 NCDMAEKVVDSSRSKSLADETGKFLFETSAKNSINVEAFISLAKDIKR---MIDTP 184
QY 203 SFQIR-----DYVESQKRSSCC 220
DB 185 NEQPVVQPGTNGANNKKACC 208

RESULT 15
ID YPT2_MAIZE STANDARD; PRT; 203 AA.
AC Q05737;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE GTP-binding protein YPTM2.
GN YPTM2.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Coleoptile;
RX MEDLINE=92115746; PubMed=1731354;
RA Palme K., Diefenthal T., Vington M., Sander C., Schell J.;
RT "Molecular cloning and structural analysis of genes from Zea mays
(L.) coding for members of the ras-related ypt gene family.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:787-791(1992).
CC -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC TRAFFIC (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: ITS EXPRESSION IS WEAK IN STEMS, HIGHER IN
CC ROOTS, LEAVES AND COLEOPTILES, BUT HIGHEST IN FLOWERS.
CC -!- PTM: THE PALMITIC ACID IS REQUIRED FOR MEMBRANE ATTACHMENT AND
CC BIOLOGICAL FUNCTION
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC -----
CC EMBL; X63278; CAA44919.1; -.
CC PIR; B38202; B38202.
CC HSP; P05713; 3RAB.
CC MaizeDB; 78605; -.
CC InterPro; IPR003579; GTPase_Rab.
CC InterPro; IPR001806; Ras_trnsfrmg.
CC InterPro; IPR005225; Small_GTP.
CC Pfam; PF00071; ras; 1.
CC PRINTS; PR00449; RASTRNSFRMG.
CC SMART; SM00175; RAB; 1.
CC TIGRFAMS; TIGR00231; small_gtp; 1.
CC GTP-binding; Lipoprotein; Prenylation; Palmitate; Protein transport.
KW NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (PROBABLE).

FT LIPID 200 200 PALMITATE (BY SIMILARITY).
FT LIPID 201 201 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 203 AA; 22475 MW; E241326E7ACD1B8A CRC64;
Query Match 43.0%; Score 495; DB 1; Length 203;
Best Local Similarity 45.1%; Pred. No. 1.2e-36;
Matches 92; Conservative 51; Mismatches 53; Indels 8; Gaps 3;

QY 23 SPSYDLTGKVMILGDTGCGKTCFLQKDGAFLSGTFIATVGIDFRNKVYTVDGVVRVKLQ 82
DB 2 NPEYDYLFKLLLLIGDSGVKSCLLLRFSDFSYLD-SYISTIGVDFKIRITVEQDGKTIKIQ 60
QY 83 IWDTAGCQERFRSVTHAYYRDAQALLLLYDITNKKSSFDNIRAMLTETHEYAQRDVMIMLG 142
DB 61 IWDTAGCQERFRITATYYRGAGILLVYDVTDEKSFGRNIRNIRTEQHATDSVNMKMLIG 120
QY 143 NKADMSERVIRSEDEGTALAREYGVFPLETSAGTKGNVELAFLAIAKELKYRAGHQ--- 198
DB 121 NKSDLTPANKVAVATETAKAFADENGIPFMETSAKNATNVQOAFMAAASIKDRMASQPAAA 180
QY 199 -ADEPSFQIRDYVESQKRSSCCS 221
DB 181 NARPATVQIRGQPVNOK--TSCCS 202

Search completed: November 17, 2002, 16:40:09
Job time : 26 sec's

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 17, 2002, 16:41:58 ; Search time 1955 Seconds
(without alignments)
1847.362 Million cell updates/sec

Title: US-09-817-199A-2
Perfect score: 1150
Sequence: 1 MTGTPGAVATRDGEAPERSP.....FQIRDYVESQKRSCCSFM 223

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09817199/runat_13112002_134902_5176/app_query.fasta_1.391
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -DOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09817199.ecgn.1.1.899.erunat_13112002_134902_5176 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1150	100.0	1014	14	BM921365 AGENCOURT
2	1137	98.9	1051	12	EG283602
3	1134	98.6	887	9	AL522282
4	1058	92.0	676	10	BB598938
5	1020	88.7	651	10	BB633978
6	1006	87.5	740	13	BI767046
7	1001	87.0	655	10	BB635649
8	856.5	74.5	749	9	AL559085
9	856	74.4	733	10	BB619625
10	852	74.1	522	13	BM151643
11	829	72.1	736	13	EG934396
12	825	71.7	676	10	BB206788
13	790	68.7	553	12	BE755280
14	763	66.3	488	13	BM149118
15	760	66.1	888	14	BQ687186
16	760	66.1	953	12	BE906081
17	754	65.6	895	14	BQ688147
18	752	65.4	579	10	BB621927
19	740	64.3	657	10	BB196489
20	739	64.3	455	10	AW464449
21	728	63.3	691	12	BG253976
22	721.5	62.7	672	13	BM492040
23	650	56.5	701	10	AW914157
24	646	56.2	708	9	AJ442405
25	622	54.1	780	12	BG784310
26	618	53.7	443	10	AW956878
27	610	53.0	380	10	BE241772
28	607	52.8	386	10	AW484587
29	588	51.1	820	13	BI916873
30	582	50.6	782	13	BI545231
31	581.5	50.6	658	10	AV697217
32	581	50.5	368	13	BM152052
33	580.5	50.5	648	10	AV697216
34	580.5	50.5	684	10	AV692701
35	580.5	50.5	697	10	AV692689
36	576.5	50.1	678	10	AV685389
37	572	49.7	668	10	AV688197
38	557.5	48.5	647	10	AV698293
39	557	48.4	678	10	AV686507
40	548	47.7	692	12	BG699266
41	535	46.5	866	9	AU079251
42	530	46.1	592	13	BM023074
43	528.5	46.0	767	12	BF608560
44	528.5	46.0	939	14	BQ947443
45	527	45.8	404	13	BM149006

ALIGNMENTS

RESULT 1
BM921365
LOCUS
DEFINITION AGENCOURT_6626159 NIH_MGC_115 Homo sapiens cdna clone IMAGE:5752779
5', mRNA sequence.
ACCESSION BM921365
VERSION BM921365.1 GI:19371744
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1014)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12787 row: i column: 04
High quality sequence stop: 735.
Location/Qualifiers

FEATURES
source

1. .1014
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5752779"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT 237 a 288 c 295 g 194 t

ORIGIN

Alignment Scores:
Pred. No.: 6,16e-139 Length: 1014
Score: 1150.00 Matches: 223
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-817-199a-2 (1-223) x BM9211365 (1-1014)

QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
Db 48 ATACGGGGACGGCCGGCGGTGGCCACCGGGATGGCGAGGCCCGCCGCGCTCCCG 107
QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
Db 108 CCCTGCAGTCCGAGCTACGACCTCAGCGGCAAGGTGATGCTTCTGGAGACACAGCGGTC 167
QY 41 GlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
Db 168 GGCACAAACATGTTCTGTGATCCAAATCAAGAGGGGGCTTCTGTCGCGAACCTTCATA 227
QY 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
Db 228 GCACCGCTGGCAGATACATTCAGGAACAAGGTGGTCACTGTGGATGGCGGTGAGTGAAG 287
QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
Db 288 CTGCAGATCTGGACACCGCTGGCGAGGACGGTTCGAGAGCGTCAACCATGCTTTATTC 347
QY 101 ArgAspAlaGlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
Db 348 AGAGATGCTCAGGCCCTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 407
QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValIleMetLeu 140
Db 408 ATCAGGGCTGGCTCACTGAGATTCATGATGATGATGATGATGATGATGATGATGATGAT 467
QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
Db 468 CTAGGCAACAAGCGGATATGACACGGAAGAGTATCCGTTCCGAACGCGAGAGACC 527

QY 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
Db 528 TTGGCCAGGAGTACGGTGTTCCTTCTGGAGACCGCCAAAGACTGCATGATGTG 587
QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
Db 588 GAGTTAGCCTTTCTGGCCATCGCAAGAACTGAAATACCGGCGCGGCATCAGGCGGAT 647
QY 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
Db 648 GAGCCAGCTTCCAGATCCGAGACTATGTAGTCCAGAGAAAGCGCTCCAGCTGCTGC 707
QY 221 SerPheMet 223
Db 708 TCCTTCATG 716

RESULT 2

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-817-199a-2 (1-223) x BG283602 (1-1051)

QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20

Db 3 ATGACGGGACCGCCGAGGCGCGGTGGCCACCGGGATGGCGAGGCCCGCCGCGCTCCCG 62

QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40

Db 108 CCCTGCAGTCCGAGCTACGACCTCAGCGGCAAGGTGATGCTTCTGGAGACACAGCGGTC 167

QY 41 GlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60

Db 168 GGCACAAACATGTTCTGTGATCCAAATCAAGAGGGGGCTTCTGTCGCGAACCTTCATA 227

QY 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80

Db 228 GCACCGCTGGCAGATACATTCAGGAACAAGGTGGTCACTGTGGATGGCGGTGAGTGAAG 287

QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100

Db 288 CTGCAGATCTGGACACCGCTGGCGAGGACGGTTCGAGAGCGTCAACCATGCTTTATTC 347

QY 101 ArgAspAlaGlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120

Db 348 AGAGATGCTCAGGCCCTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 407

QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValIleMetLeu 140

Db 408 ATCAGGGCTGGCTCACTGAGATTCATGATGATGATGATGATGATGATGATGATGATGAT 467

QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160

Db 468 CTAGGCAACAAGCGGATATGACACGGAAGAGTATCCGTTCCGAACGCGAGAGACC 527

QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20

Db 3 ATGACGGGACCGCCGAGGCGCGGTGGCCACCGGGATGGCGAGGCCCGCCGCGCTCCCG 62

QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40

Db 108 CCCTGCAGTCCGAGCTACGACCTCAGCGGCAAGGTGATGCTTCTGGAGACACAGCGGTC 167

QY 41 GlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60

Db 168 GGCACAAACATGTTCTGTGATCCAAATCAAGAGGGGGCTTCTGTCGCGAACCTTCATA 227

QY 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80

Db 228 GCACCGCTGGCAGATACATTCAGGAACAAGGTGGTCACTGTGGATGGCGGTGAGTGAAG 287

QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100

Db 288 CTGCAGATCTGGACACCGCTGGCGAGGACGGTTCGAGAGCGTCAACCATGCTTTATTC 347

QY 101 ArgAspAlaGlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120

Db 348 AGAGATGCTCAGGCCCTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 407

QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValIleMetLeu 140

Db 408 ATCAGGGCTGGCTCACTGAGATTCATGATGATGATGATGATGATGATGATGATGATGAT 467

QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160

Db 468 CTAGGCAACAAGCGGATATGACACGGAAGAGTATCCGTTCCGAACGCGAGAGACC 527

QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20

Db 3 ATGACGGGACCGCCGAGGCGCGGTGGCCACCGGGATGGCGAGGCCCGCCGCGCTCCCG 62


```
QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
|||||
Db 63 CCTCAGTCCGAGTACAGCTCAGGCGCAGGTGATGCTCTCTGGGACACAGCGCTC 122
QY 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
|||||
Db 123 GGCACAAAGATGTTCTCTGATCCAAATCAAGACGCGGCTTCTCTCGGAACCTTCATA 182
QY 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
|||||
Db 183 GCCACCGTCGGCATAGACTTCAGAACCAAGGTGGTGACTGTGGTGGGTGAGATGAAG 242
QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
|||||
Db 243 CTGCAGATCTGGCAGCACCGCTGGCGAGCAACGGTTCGGAAGGCTCACCCATCTATTAC 302
QY 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
|||||
Db 303 AGAGATGCTCAGCGCTTGTCTGTGTATGACATCACCAACAAATCTCTTTTCGACAC 362
QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
|||||
Db 363 ATCAGGGCTGCTCCTCAGTATGATGATGATGATGATGATGATGATGATGATGATG 422
QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
|||||
Db 423 CTAGGCAACAAGCGGATATGAGCAGCAAGAGTGTATCCGTTCGGAAGACGAGAC 482
QY 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
|||||
Db 483 TTGGCAGGGAGTAGGGTGTCTCTCTGGAGACGAGCGCAAGACTGGCATGAATGTG 542
QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
|||||
Db 543 GAGTTAGCTTCTGGCCATCCCAAGGAACCTGAATACCGGCGGGCATCAGCGGAT 602
QY 201 GluProSerPheGlnIleArgAspTyrValGluSerGln-LysLysArgSerSerCysCys 220
|||||
Db 603 GAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAGAGCGCTCCAGCTGCTG 662
QY 220 sSerPheMet 223
|||||
Db 663 CTCCTTCATG 672

RESULT 3
AL522282 LOCUS AL522282 887 bp mRNA linear EST 13-FEB-2001
DEFINITION AL522282 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB008YF08 5
prime, mRNA sequence.
ACCESSION AL522282
VERSION AL522282.1 GI:12785775
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 887)
AUTHORS Li.W.B., Gruber.C., Jessee.J. and Polayes.D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 887
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB008YF08"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"

FEATURES
source
```

```
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 203 a 256 c 257 g 168 t
ORIGIN

Alignment Scores:
Pred. No.: 6 08e-137 Length: 887
Score: 1134.00 Matches: 220
Percent Similarity: 99.10% Conservative: 1
Best Local Similarity: 98.65% Mismatches: 2
Query Match: 98.61% Indels: 0
Db: 9 Gaps: 0

US-09-817-199a-2 (1-223) x AL522282 (1-887)
QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
|||||
Db 28 ATGACGGGCGACGCCAGGCGCGCTTGCACCCGGGATGGCGAGCCGCCGAGCGCTCCCG 87
QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
|||||
Db 88 CCTCGCACTCCGAGCTACGACCTCAGGGCAAGGTGATCTCTGGGAGACACAGCGCTC 147
QY 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
|||||
Db 148 GCAAAACATGTTCTCTGATCCAAATTCAAAGACGGGCGCTTCTCTCCGGAACCTTCATA 207
QY 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
|||||
Db 208 GCCACCGTCGGCATAGACTTCAGGAACAAGTGGTGACTGTGGATGGCTGAAAGTGAAG 267
QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
|||||
Db 268 CTGCAGATCTGGGACACCGCTGGCGAGGAACGGTTCGGAAGCGTCAACCATGCTATTAC 327
QY 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
|||||
Db 328 AGARATGCTCAGCGCTTGTCTGTGTATGACATCACCAACAAATCTCTTCTCGACAC 387
QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
|||||
Db 388 ATCAGGGCTGGCTCACTGAGATTCTATGATGATGCCAGAGGACGCTGGTATCATGCTG 447
QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
|||||
Db 448 CTAGGCAACAAGCGGATATGAGCAGCAAAAGTGTATCCGTTCGGAAGCGAGAGACC 507
QY 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
|||||
Db 508 TTGGCCAGGGAGTACGGTGTCTCTCTCTGGAGACCGAGCCCAAGACTGGCATGAATGTG 567
QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
|||||
Db 568 GAGTTAGCTTCTTCTGGCCATCCCAAGGAACCTGAATACCGGCGGCGCATCAGCGGAT 627
QY 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysArgSerSerCysCys 220
|||||
Db 628 GAGCCCAAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAAAGCGCTCCAGCTGCTGC 687
QY 221 SerPheMet 223
|||||
Db 688 TCCTTCATG 696

RESULT 4
```

BB598938 676 bp mRNA linear EST 26-OCT-2001
LOCUS BB598938 RIKEN full-length enriched, adult pancreas islet cells Mus musculus cDNA clone C820003E14 5', mRNA sequence.
ACCSSION VERSION BB598938
KEYWORDS house mouse.
SOURCE ORGANISM Mus musculus
REFERENCE AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 676)
TITLE JOURNAL COMMENT Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,Y., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.
FEATURES RIKEN Mouse ESTs (Arakawa,T., et al. 2001) Unpublished (2001)
On Dec 1, 2000 this sequence version replaced gi:11507539.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1637-1630 (2000)
wagi,K., Fujiwara,K., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanishi,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaoka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata.K. and Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Location/Qualifiers
1..676
/oranism="Mus musculus"
/db_xref="taxon:10090"
/clones="C820003E14"
/clone_lib="RIKEN full-length enriched, adult pancreas islet cells"
/tissue_type="pancreas"
/cell_type="islet cells"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: Sali; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia"

LOCUS	BI767046	740 bp	linear	EST 25-SEP-2001
DEFINITION	6030545267f1 NIH MGC 122 Homo sapiens cDNA clone IMAGE:5203600 5'			

QY IOI ARGASPATAGINATALEULEULEULTYFASPIETHTHFAASHLYSSERSEPHEASPN 120

US-09-81

US-09-817-199A-2 (1-223) x AL559085 (1-749)


```
Db 3 ACTGGCACACTAGGAGCTGCTACCGCTGGGGATGCGGAGCGCCCTCAGCGCTCCCGCCGC 62
Qy 22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41
Db 63 TTCACCCGCACTAGATCTCACCGGCAAGGTGATGCTCTTGGAGACTCGGCGCTGGC 122
Qy 42 LysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAla 61
Db 123 AAACCTGCTTCTCTGATCCAAATTCAAAGACGGGGCTCTCTCCGGAACCTTCATAGCC 182
Qy 62 ThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeu 81
Db 183 ACCGTCGCGCATAGACTTCAGGAATAAGTGTGACATGGATGGTCCGCGGATGAAGCTT 242
Qy 82 GlnIleTyrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArg 101
Db 243 CAGATCTGGGACACATGACAGACAGAGCGCTTCCGCGATGTGACCATGCTTATTACCGA 302
Qy 102 AspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIle 121
Db 303 GATGCTCAAGCTTGTCTCTGTGTATGACATCACCAACCAATCTCTTTTTCACAACATA 362
Qy 122 ArgAlaTyrLeuThr-GluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 141
Db 363 AGGCGCTGGCTTACAAAAGATTTCATGAAGTTCGCCAGAGACGCTGGTGATTATGCTTTT 422
Qy 141 uGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrIle 161
Db 423 AGGCAC-AAGGCGGATGTAAGCAGCGAAAGGGTGATCGCTTGAAGAAGGAACACT 481
Qy 161 uAlaArgGluTyrGlyValProPheLeuGlu-ThrSerAlaLysThrGlyMetAsnValG 181
Db 482 GCGCAGGAATATGCTGGCTCTTCATGAGAACCGTGGCCAAAGACTGGCTTGAACGTGG 541
Qy 181 uLeuAla-PheLeuAlaIleAlaLysGlu-LeuLysTyrArgAla-GlyHisGlnAlaA 200
Db 542 AGTGGCCCTTCTTGGCAATTCGCCAAGGAACTTAAATACCGTGAAGGAGCGACTCTG 601
Qy 200 spGluProSerPheGlnIleArgAspTyrValGlu-SerGlnLysLysArgSerSerCys 219
Db 602 ATGACCCCACTTTCAAAACCGAGATATGTGGAGTCCCAAAAACCGCTCCCAACTGG 561
Qy 220 Cys 220
Db 662 TGC 664

RESULT 10
BM151643
LOCUS
DEFINITION
522 bp mRNA linear EST 30-NOV-2001
TCBAP1E10302 Pediatric pre-B cell acute lymphoblastic leukemia
Baylor-HGSC project-TCBA Homo sapiens cDNA clone TCBAP1030, mRNA
sequence.
ACCESSION BM151643
VERSION BM151643.1 GI:17175247
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (Bases 1 to 522)
Weil,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R. Jr.,
Gunarathne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
Pediatric Leukemia cDNA Sequencing Project (2001)
JOURNAL Unpublished (2001)
COMMENT
Contact: Dr. Judith F. Margolin
Texas Children's Cancer Center and Human Genome Sequencing Center
at Baylor College of Medicine
1102 Bates, MC3-3320 Houston, TX 77030, USA
Tel: 832-824-4536
Fax: 832-825-4038
Email: clones@tccc.org
Seq primer: M13 primer.
Location/Qualifiers
```

```
source
1..522
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TCBAP1030"
/clone_lib="Pediatric pre-B cell acute lymphoblastic
leukemia Baylor-HGSC project-TCBA"
/sex="male"
/tissue_type="leukopheresis"
/cell_type="pre-B cell"
/dev_stage="Pediatric 2 years"
/lab_host="DH10B"
/Note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGCGCGGAGGAG(T)VN
3': V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
was primed with a BamHI-dC primer
[5'AGAGAGCTCGGATCCGCCCAATAAATAAT(C)3'].
Double-stranded cDNA was then digested with BamHI and XhoI
and directionally cloned into the BamHI and SalI sites of
lambda pSB vector. Library went through one round of
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Carninci P. Westover A, Nishiyama Y, Ohsumi T,
Itoh M, Nagaoka S, Sasakin, Okazaki Y, Muramatsu M,
Schneider C, Hayashizaki Y, High efficiency selection of
full-length cDNA by improved biotinylated cap trapper.,
DNA Res 4: 1, 61-6, Feb 28, 1997)"
BASE COUNT 119 a 146 c 158 g 99 t
ORIGIN
Alignment Scores:
Pred. No.: 1.6e-100 Length: 522
Score: 852.00 Matches: 165
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 74.09% Indels: 0
DB: 13 Gaps: 0
US-09-817-199A-2 (1-223) x BM151643 (1-522)
Qy 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
Db 19 ATGAGGGGACCGCCAGCGCGCTTGCACCCCGGATGGGAGCGCCCTTCCTGTCGGAACTTCATA 78
Qy 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
Db 79 CCTGCAGTCCGAGCTACGACCTCAGGGGCAAGGTGATGCTTCTGGAGACACACAGCGCTC 138
Qy 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
Db 139 GGCAAAACATGTTTCTGTATCCAAATTCAAAGACGGGGCTTCCTGTCGGAACTTCATA 198
Qy 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
Db 199 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGAAG 258
Qy 81 LeuGlnIleTyrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
Db 259 CTGCAGATCTGGACACCCGCTGGGAGGAACCGTTCGGAAGCGCTCACCATGCTTATTAC 318
Qy 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
Db 319 AGAGATGCTCAGGCGCTTGTCTGTGTATGACATCACCAACAAATCTTCTTTTCGACAAC 378
Qy 121 IleArgAlaTyrLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
Db 379 ATCAGGGCTGGCTCACTGAGATTATGAGTATGCCAGAGGACGTGGTGATCATGCTG 438
Qy 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
Db 439 CTAGGCACAAAGCGGATATGAGCAGCGAAAGAGTATCGTTCGCAAGACGAGAGACC 498
Qy 161 LeuAlaArgGluTyr 165
|||||
```


Db 363 ACTGGCATGAACGTGGAGTTGGCCTTCTTGGCAATTGCCAGGAACCTGAATACCGTGCA 42

Qy 196 GlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSerGlnLys 21

Qy 21 p 1

Qy 21 p 1

Db	131	CCCTCGGGCCGAGCCAGCATCTCAGGGCAAGGTGATGCTTCTCGGGAGACTCGGGCGCTC	190
Qy	41	GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle	60
Db	191	GGCAAAACCTGTTTCTCATCGATTACAGACAGGGGCTTCTCTCGGGACCTTCATA	250
Qy	61	AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValargVallys	80
Db	251	GCCACCGTCGGCATAGACTTCAGGAACAAATGGTGACCGTGGATGGTGTGAGAGTGAAA	310
Qy	81	LeuGlnIleTipAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaIleTyr	100
Db	311	CTGCAGATCTGGGACACGCGAGGGCAGCGGTCCGACGTGTACCCATGCTTATTAC	370
Qy	101	ArgaspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn	120
Db	371	CGAGACGCCAGGGCTTGTCTCTGCTGTCAGACATCACCAAAATCTTCCTTTGACAAC	430
Qy	121	IleArgAlaIleLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu	140
Db	431	ATCGGGCTTGGCTCAGTACAGATTCACGATGACGCCAGAGGATGTGGTATCATGCTT	490
Qy	141	LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr	160
Db	491	CTGGGCACAAAGGAGATGTGAGCACTGAAAGGCTGATCCGCTCAGAGCGGAGAGATG	550
Qy	161	Leu 161	
Db	551	CTG 553	
RESULT	14		
LOCUS	BM149118		
DEFINITION	TCAAP2E6371 Pediatric acute myelogenous leukemia cell (FAB M1) Baylor-HGSC project-TCAA Homo sapiens cDNA clone TCAAP6371, mRNA sequence.		
ACCESSION	BM149118		
VERSION	BM149118.1	GI:17170448	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Wei, X., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman, F.R. Jr., Gunaratne, P.H., Muzny, D., Bouck, J., Gibbs, R.A., and Margolin, J.F.		
TITLE	Pediatric Leukemia cDNA Sequencing Project (2001)		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Dr. Judith F. Margolin Texas Children's Cancer Center and Human Genome Sequencing Center at Baylor College of Medicine 1102 Bates, MC3-3320 Houston, TX 77030, USA Tel: 832-824-4536 Fax: 832-825-4038 Email: clones@ccc.org Seq primer: M13 primer		
FEATURES	Location/Qualifiers		
source	1. .488 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="TCAAP6371" /clone_lib="Pediatric acute myelogenous leukemia cell (FAB M1) Baylor-HGSC project-TCAA" /sex="male" /tissue_type="leukopheresis" /cell_type="myeloid cell" /dev_stage="pediatric 6 years" /lab_host="DH10B" /note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI; First strand cDNA was primed with an anchored XhoI-oligo(dr) primer [5'GGAGACTCGAGCGGCCGAGGAG(T)VN 3'; V=A,C,G; N=A,C,G,T] and then dg tailed. Second strand was primed with a BamHI-dC primer		

Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LCM2362 row: a column: 14
 High quality sequence stop: 662.

FEATURES
 source
 Location/Qualifiers
 1..888
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6206869"
 /clone_lib="NIH_MGC_110"
 /tissue_type="ductal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAGCAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

BASE COUNT 172 a 279 c 289 g 148 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3.77e-88 Length: 888
 Score: 760.00 Matches: 149
 Percent Similarity: 76.62% Conservative: 28
 Best Local Similarity: 64.50% Mismatches: 30
 Query Match: 66.09% Indels: 24
 14 Gaps: 3

US-09-817-199A-2 (1-223) x B0687186 (1-888)

Qy 4 ThrProGlyAla-----ValAlaThrArgAspGlyGluAlaProGluArgSer----- 19
 Db 81 ACCCCCGCTGCCCTCCACGCTGCCACCGCCACACGGGGCCGACCGCGCTCCGGGACT 140
 Qy 20 -----ProCysSerPro----- 24
 Db 141 GCGCTTCCGGCCCGACGCGCGCCCAACGGCCCTTCACGCCCGCGCGCTCGCTT 200
 Qy 25 -----SerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGly 39
 Db 201 GCGCGCGGTGTCGACTTCTACGACGTCGCCTTCAAGGTCATGCTGGTGGGGACTCGGT 260
 Qy 40 ValGlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPhe 59
 Db 261 GTGGGAGACCTGCTGCTGTGGGATTCAGGATGGTCTTCTGGCGGGACCTTC 320
 Qy 60 IleAlaThrValIleAspPheArgAsnLysValThrValAspGlyValArgVal 79
 Db 321 ATCTCCACGCTAGGCATTGACTTCCGGAAACAAGTCTGGACGTGGTGTGAAGGTG 380
 Qy 80 LysLeuGlnIleThrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyr 99
 Db 381 AAGCTGCAGATGTGGGACACAGCTGCTCAGGAGCGGTTCCGCAGTCTTACCCATGCCCTAC 440
 Qy 100 TyrArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAsp 119
 Db 441 TACCGGGATGCTCATGCTCTGCTGCTCTACGATGTACCAACAAGGCCCTCCTTTGAC 500
 Qy 120 AsnIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMet 139
 Db 501 AACATCCAGGCTGGCTGACCGAGATCCACGAGTACGCCACACGACGCTGGCGCTCATG 560
 Qy 140 LeuLeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGlu 159
 Db 561 CTGCTGGGGAACAAGGTGGACTCTGCCCATGAGCGTGTGGTGAAGAGGGAGGACGGGAG 620

Qy 160 ThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsn 179
 Db 621 AAGCTGGCCCAAGGAGTATGGACTGCCCTTCATGGAGACCGCCCAAGACGGGCTCAAC 680
 Qy 180 ValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAla 199
 Db 681 GTGGACTTGGCCTTCACAGCCATAGCAAGAGGAGTGAAGCGCTCCCTGAAGGCTCCC 740
 Qy 200 AspGluProSerPheGlnIleArgAspTyrVal 210
 Db 741 ACGAGACCGCGCTTCGGCTGCATGATTACGTT 773

Search completed: November 17, 2002, 18:05:43
 Job time : 1959 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2002, 16:01:33 ; Search time 81 Seconds
(without alignments)
567.266 Million cell updates/sec

Title: US-09-817-199A-2
Perfect score: 1150
Sequence: 1 MTGTPGAVATRDGEAPERSP.....FQIRDYVESQKKRSCGSEFM 223

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.rodent.*
- 12: sp.virus.*
- 13: sp.vertibrate.*
- 14: sp.unclassified.*
- 15: sp.virus.*
- 16: sp.bacteriap.*
- 17: sp.archeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	ID	Description
1	697	60.6	666	5	Q9VP48	Q9VP48 drosophila
2	648.5	56.4	388	5	Q9SR32	Q9SR32 drosophila
3	533	46.3	196	5	O02046	O02046 caenorhabdi
4	521.5	45.3	207	11	O8VCF6	O8VCF6 mus musc
5	514.5	44.7	202	10	Q9SXT5	Q9SXT5 cicier ariet
6	513.5	44.7	202	10	O8L153	O8L153 pisum sativ
7	510	44.3	216	10	O49844	O49844 daucus caro
8	509.5	44.3	214	10	O40218	O40218 lotus japon
9	509	44.3	203	10	Q949E2	Q949E2 oryza sativ
10	509	44.3	216	10	O40215	O40215 lotus japon
11	509	44.3	218	10	Q9SF91	Q9SF91 arabidopsis
12	508.5	44.2	204	5	O15971	O15971 drosophila
13	507.5	44.1	202	10	O40203	O40203 lotus japon
14	507	44.1	216	10	Q9LZD4	Q9LZD4 arabidopsis
15	507	44.1	216	10	O8VWF9	O8VWF9 nicotiana t
16	506	44.0	203	10	Q9M7P5	Q9M7P5 capsicum an

17	506	44.0	216	10	Q40177	Q40177 lycopersico
18	505.5	44.0	201	5	Q94148	Q94148 caenorhabdi
19	505	43.9	203	10	O24112	O24112 nicotiana p
20	505	43.9	216	10	O8W3J4	O8W3J4 nicotiana t
21	504	43.8	216	10	O8W3J3	O8W3J3 nicotiana t
22	503	43.7	203	10	O41340	O41340 lycopersico
23	502	43.7	216	10	O9FJF1	O9FJF1 arabidopsis
24	501.5	43.6	203	10	O8RU63	O8RU63 oryza sativ
25	501	43.6	203	10	O940Z7	O940Z7 arabidopsis
26	500	43.5	216	10	O24466	O24466 arabidopsis
27	499.5	43.4	216	10	O8W3J2	O8W3J2 nicotiana t
28	499	43.4	212	10	O40217	O40217 lotus japon
29	498	43.3	203	10	O41338	O41338 lycopersico
30	498	43.3	216	10	O9SW8	O9SW8 arabidopsis
31	497.5	43.3	202	10	O9FPJ4	O9FPJ4 arabidopsis
32	497.5	43.3	202	10	O8L155	O8L155 pisum sativ
33	497.5	43.3	215	10	O41023	O41023 pisum sativ
34	496.5	43.2	202	10	O9SEH3	O9SEH3 arabidopsis
35	495	43.0	215	10	O40219	O40219 lotus japon
36	495	43.0	216	10	O41024	O41024 pisum sativ
37	490.5	42.7	215	10	O41022	O41022 pisum sativ
38	489	42.5	222	10	O96362	O96362 brassica ca
39	488	42.4	203	10	O8L154	O8L154 pisum sativ
40	488	42.4	203	10	O40569	O40569 nicotiana t
41	487.5	42.4	202	10	O8W4S8	O8W4S8 arabidopsis
42	487.5	42.4	207	5	O18338	O18338 drosophila
43	487	42.3	200	4	O9H0T3	O9H0T3 homo sapien
44	486	42.3	202	10	O9ZRH6	O9ZRH6 petunia hyb
45	484.5	42.1	215	10	O41061	O41061 pisum sativ

ALIGNMENTS

RESULT 1

Q9VP48	PRELIMINARY;	PRT;	666 AA.
ID	Q9VP48;		
AC	Q9VP48;		
DT	01-MAY-2000 (Tremblrel. 13, Created)		
DT	01-MAY-2000 (Tremblrel. 13, Last sequence update)		
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)		
DE	CG7605 protein.		
GN	CG7605.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RX	SEQUENCE FROM N.A.		
RP	STRAIN=BERKELEY.		
RC	MEDLINE=20190006; PubMed=10731132;		
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	April J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,		
RA	Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RA	Fosler K., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,		
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,		
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,		
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		

85 DTAGQERFSVTHAYYRDAQAALLLYDITNKSSFDNIRAWLWTEIHEYAQRDVVVIMLLGNK 14 QY
 :|||:||| || | || | :||| ||| :||| ||| :||| ||| :||| ||| :||| |||
 Db 63 DTAGQERFRTTAYYRGAMGIVLVVDITNEKSDFINRWIRNIEHGSADYEKMILGNK 122
 :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
 QY 145 ADMSESVIRSDEGTALREYGVPLETSAKTGMMVELAFLAIAKEELKYR-----AGHOA 199
 Db 123 CDVNDRQVKSGERGEKLADYGIKFMTESAKANINVENAFFTLARDIKAKMKQLLEGNSP 182
 : ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
 QY 200 DEPSFIQIRDYVESOKRRSCC-CSFM 223 PRELIMINARY; PRT; 202 AA.
 Db 183 QGSSHGVKITVEQKRFTSFRCSSL 207
 : ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

RESULT 5
 Q9SXT5
 ID AC Q9SXT5 PRELIMINARY; PRT; 202 AA.
 DB DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-JUN-2002 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Rab-type small GTP-binding protein.
 OS Cicer arietinum (Chickpea) (Garbanzo).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae; Cicer.
 OX NCBI_TaxID=3827;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=ILC3276; TISSUE=LEAF;
 RA Ichinose Y., Tiemann K., Schwenger-Erger C., Toyoda K., Hein F.,
 RA Hanselle T., Barz W.;
 RT "Genes Expressed in Ascochyta rabiei-Inoculated-Chickpea Plants and
 RT Elicited Cell Cultures as Detected by Differential cDNA-
 RT Hybridization.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DDAJ databases.
 CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 EMBL; AB024994; BAA76422.1; -.
 DR HSP; P01112; IPLK.
 DR InterPro; IPRO03579; GTPase_Rab.
 DR InterPro; IPRO01806; Ras_trnsmfrng.
 DR InterPro; IPRO02078; Sig54_interact.
 DR InterPro; IPRO05225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTNRNFRMNG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRFAMS; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
 KW GTP-binding; Lipoprotein.
 SQ SEQUENCE 202 AA; 22487 MW; 8CC4FCBAAC0C8CCE CRC64;

Query Match 44.7%; Score 514.5; DB 10; Length 202;
 Best Local Similarity 47.3%; Pred. No. 1.7e-39;
 Matches 96; Conservative 47; Mismatches 53; Indels 7; Gaps

QY 23 SPFSYDLTKVMILLGDVGKTCTFLIQDKGAFLSGPTFAVTGIDFNKVTVYDGVVRVKLO 82
 Db :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
 2 NPEDYLFLKLLIGDSGVGKSCLLRLRFADDSYLD-SYSTIGVFDIRVVEQDGKTIKLQ 60
 QY 83 INDTAQGERFSRVTHAYYRDAQAALLLYDITNKSSFDNTRAWLTEIHEFYAQRDVVVIMLG 142
 Db :|||:||| || | || | :||| ||| :||| ||| :||| ||| :||| ||| :||| |||
 Db 61 INDTAQGERFRITSYYRGAGIIIVYDVTQDESNNYKOWNLIEDRYASENVNKKLLVG 120
 : ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
 QY 143 NKADMSSERVIRSEDGETALREYGVPFLETSAKTGMNVELAF LAIAKEELKYRAGHQ---- 198
 Db :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
 Db 121 NKCGLANRVVSSETAKFADEIGIPFMTESAKNATNV EQAFMAMAEEIKRNMASOPANN 180
 QY 199 ADEPFQIRDYVESOKRRSCCS 221
 Db : ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
 Db 181 ARPPTVIIRGPMPNQK--SGCCS 201

RESULT 6
 Q08153

```
ID O08153 PRELIMINARY; PRT; 202 AA.
AC Q08153;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE GTP-binding protein.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RX MEDLINE=94290824; PubMed=8019783;
RA Nagano Y., Murai N., Matsuno R., Sasaki Y.;
RT "Isolation and characterization of cDNAs that encode eleven small GTP-
binding protein from Pisum sativum.";
RL Plant Cell Physiol. 34:447-455(1993).
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL; D12548; BAA02116.1; -.
DR HSSP; P05713; 3RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsfrmng.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; SM00175; RAB; 1.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 202 AA; 22498 MW; 1C9304144B5D95B3 CRC64;

Query Match 44.7%; Score 513.5; DB 10; Length 202;
Best Local Similarity 47.3%; Pred. No. 2.1e-39;
Matches 96; Conservative 47; Mismatches 53; Indels 7; Gaps 3;

QY 23 SPYDLTGKVMLLGDTGCGKTCFLIQKDGAFSLGTFIATVGIDFRNKVVTVGVRVKLQ 82
DB 2 NPEYDLFLKLLIGDSGVKSCLLRFADDSYLD-SYISTIGVDFKIRTVEQDGKTIKQLQ 60
QY 83 IWDTAGQERFSVTHAYRDAQAALLLYDITNKSSFDNIRAWLTHEIHEYAQRDVIMLIG 142
DB 61 IWDTAGQERFTTSSYIRGAHGIIVYDVTDQESFNKQWLNEDRYASENVNKLIVG 120
QY 143 NKADMSSERVIRSEGETLAREYGVPELTSAGTKMNVLEAFIAIAKELKYRAGHO--- 198
DB 121 NKSLADNKVVSSETAKAFADGIGIPMETSAKNANNVVEQAFMAAAEIKRMASQPSNN 180
QY 199 ADEPFSFOIRDYVESQKRSSCCS 221
DB 181 ARPPTVQIRGQPVNQK--SCCS 201

RESULT 7
O49844 PRELIMINARY; PRT; 216 AA.
AC O49844;
DT 01-JUN-1998 (TEMBLrel. 06, Created)
DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Small GTP-binding protein.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LUNGA DI AMSTERDAM;
RA Albani D., Bernacchia G.;

*Isolation of a cDNA encoding a rab-related small GTP-binding protein
(Accession No. AJ001367) from carrot cells (PGR97-185).";
Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
-!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL; AJ001367; CAA04701.1; -.
DR HSSP; P05713; 3RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsfrmng.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 216 AA; 23955 MW; 3277D1EEDE77175B CRC64;

Query Match 44.3%; Score 510; DB 10; Length 216;
Best Local Similarity 46.5%; Pred. No. 5e-39;
Matches 99; Conservative 47; Mismatches 55; Indels 12; Gaps 5;

QY 19 SPCSP--SYDLTGKVMLLGDTGCGKTCFLIQKDGAFSLGTFIATVGIDFRNKVVTVDG 76
DB 3 APPARADYDYLIKLLIGDSGVKSCLLRFDSGSFTT-SFTTIGIDFKIRTIEMDG 61
QY 77 VRVKLIWDTAGQERFSVTHAYRDAQAALLLYDITNKSSFDNIRAWLTHEIHEYAQRDV 136
DB 62 KRIKLIWDTAGQERFTTAYIRGAMGILLVYDVDESSFNIRNIRNIHQASDNV 121
QY 137 VMILGNKADM--SSERVIRSEGETLAREYGVPELTSAGTKMNVLEAFIAIAKELKYRA 195
DB 122 NKILVGNKADMDESKRAVPTSKQALADEYGIKFFFEASAKTNMNVVEVFFSIADIKQRL 181
QY 196 GH---QADEPFSFOIRDY----VESQKRSSCC 220
DB 182 AETDSKTEPQTIKINQOEOGAGTSAASQKSACC 214

RESULT 8
O40218 PRELIMINARY; PRT; 214 AA.
AC O40218;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE RAB8D.
GN RAB8D.
OS Lotus japonicus.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.
OX NCBI_TaxID=34305;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT NODULES;
RX MEDLINE=97231679; PubMed=9076991;
RA Borg S., Brandstrup B., Jensen T.J., Poulsen C.;
RT "Identification of new protein species among 33 different small GTP-
binding proteins encoded by cDNAs from Lotus japonicus, and expression
of corresponding mRNAs in developing root nodules.";
RL Plant J. 11:237-250(1997).
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL; Z73947; CAA98175.1; -.
DR HSSP; P05713; 3RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsfrmng.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
```



```
DR PROSITE: PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 214 AA; 23627 MW; 5E1A6B83505E50D7 CRC64;

Query Match 44.3%; Score 509.5; DB 10; Length 214;
Best Local Similarity 47.6%; Pred. No. 5.4e-39;
Matches 101; Conservative 45; Mismatches 55; Indels 11; Gaps 5;

QY 19 SPP--CSPSYDITGKVMVLGDTGVGKTCFLIOFKDGAFLSGTFTATVGVGIDFRNKVVTVDG 76
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 3 APPARADYDYLLKLLIGDGVGKSCLLRFSGDSFTT-SFTITIGIDFKIRTIELDG 61

QY 77 VYVKLIQWDTAGQERFSVTHAYYRDAQAALLLYDITNKSFDNIRAWLTETHEIYAQRDV 136
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 62 KRVKLIQWDTAGQERFTITAYYRGAMGILLVYDVTDEAFNINRWLRNIEQHASDNV 121

QY 137 VIMLLGNKADM--SSERVISEDETALAREYGVFPFLETSAKTGMNVELAFATAKELKYR- 194
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 122 NKILVGNKADMESKRAVPTSGQALADYGIKFTSAKTNUNVEVFFSIARDIKQRL 181

QY 195 --AGHQADEPSQIRD----YVESQKKRSSCC 220
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 182 ADTDHKAETTLKINQDSAAAGERAANKSSCC 213

RESULT 9
Q949E2 PRELIMINARY; PRT; 203 AA.
AC Q949E2
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Putative GTP-binding protein.
GN W455ERIPDK.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21329048; PubMed=11433398;
RA Mayer K., Murphy G., Tarchini R., Wambutt R., Volckaert G., Pohl T.,
RA Duesterhoeft A., Stiekema W., Entian K.D., Terryn N., Lemcke K.,
RA Haase D., Hall C.R., van Dodeweerd A.M., Tingey S.V., Mewes H.W.,
RA Bevan M., Bancroft I.;
RT "Conservation of microstructure between a sequenced region of the
RT genome of rice and multiple segments of the genome of Arabidopsis
RT thaliana.";
RL Genome Res. 11:1167-1174(2001).
DR EMBL; AJ307662; CAC39050.1;
DR InterPro; IPR001806; Ras_trnstrfmng.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR TIGRfams; TIGR00231; small_gtp; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding.
SQ SEQUENCE 203 AA; 22504 MW; B59DE515C09C5A82 CRC64;

Query Match 44.3%; Score 509; DB 10; Length 203;
Best Local Similarity 46.6%; Pred. No. 5.6e-39;
Matches 95; Conservative 49; Mismatches 55; Indels 8; Gaps 3;

QY 23 SPYSDLTGKVMVLGDTGVGKTCFLIOFKDGAFLSGTFTATVGVGIDFRNKVVTVDGVRVKLQ 82
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 2 NPEYDYLKLLIGDGVGKSCLLRFDSDYLD-SYISTIGVDFKIRTVQDGKTIKIQ 60

QY 83 IWDTAGQERFSVTHAYYRDAQAALLLYDITNKSFDNIRAWLTETHEIYAQRDVVIMLIG 142
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 61 IWDTAGQERFTITSSYRGAGIIVYDVTDEAFNINRWLRNIEQHASDNV 120

QY 143 NKADMSSERVISEDETALAREYGVFPFLETSAKTGMNVELAFATAKELKYRAGHQ---- 198
```

```
Db 121 NKSDLTANKVSVSETAKAFADENGIPFMETSAKNATNVQEAFAAASIKDRMASQAAA 180
QY 199 -ADEPSFQIRDYVESQKKRSSCCS 221
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 181 NARPPTVQIRGQPVNOK--TSCCS 202

RESULT 10
Q40215 PRELIMINARY; PRT; 216 AA.
AC Q40215
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE RAB8A.
GN RAB8A.
OS Lotus japonicus.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.
OX NCBI_TaxID=34305;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT NODULES;
RX MEDLINE=97231679; PubMed=9076991;
RA Borg S., Brahdstrup B., Jensen T.J., Poulsen C.;
RT "Identification of new protein species among 33 different small GTP-
RT binding proteins encoded by cDNAs from Lotus japonicus, and expression
RT of corresponding mRNAs in developing root nodules.";
RL Plant J. 11:237-250(1997).
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL; Z73944; CAA98172.1;
DR HSP; P05713; RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnstrfmng.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRfams; TIGR00231; small_gtp; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 216 AA; 23776 MW; 569926CCA8D1B489 CRC64;

Query Match 44.3%; Score 509; DB 10; Length 216;
Best Local Similarity 45.2%; Pred. No. 6.1e-39;
Matches 103; Conservative 46; Mismatches 57; Indels 22; Gaps 5;

QY 1 MTGTPGAVATRDGEAPERSPPCSPSYDITGKVMVLGDTGVGKTCFLIOFKDGAFLSGTFT 60
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 1 MAGAPAPARAD-----YDYLKLLIGDGVGKSCLLRFSGDSFTT-SFT 46

QY 61 ATVGIDFRNKVVTVDGVRVKLQIWDTAGQERFSVTHAYYRDAQAALLLYDITNKSFDN 120
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 47 TTIGIDFKIRTIELDGKRIKLIQWDTAGQERFTITAYYRGAMGILLVYDVTDEASENN 106

QY 121 IRAWLTETHEIYAQRDVVIMLLGNKADM--SSERVISEDETALAREYGVFPFLETSAKTMN 179
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 107 IKNWIRNIEQHASDNVKNILVGNKADMESKRAVPTSGQALADYGIKFTSAKTNLN 166

QY 180 VELAFATAKELKYR---AGHQADEPSFQIRDYVES---QKKRSSCC 220
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 167 VEEVFFSIARDIKQRLADTDSRAEPQTIQINQPDASASGGQAQKSCC 214

RESULT 11
Q9SF91 PRELIMINARY; PRT; 218 AA.
AC Q9SF91
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
```


QY 25 SYDLTGKVMLLGDTGVGKTCFLIOFKDGAFLSGTFTIATVGDIDFRNKVVTVGVVRVKLQIW 84
 DB 5 TYDLLFKLLIGDSCGVGKTCILFRSDDAFTS-TFISTIGIDFKIKTVELRGKIKLQIW 63
 QY 85 DTAGQERFRSVTHAYYRDAQAALLLLYDITNKSSFNIRAWLTIEHYAQRDQVVMILGNK 144
 DB 64 DTAGQERFRFTTSYRGAMGIMLYDITNKSSFNIRAWLTIEHYAQRDQVVMILGNK 123
 QY 145 ADMSSERVISGDTGLAREGVPPLETSAKTMVLAFLAIAKE-LKYRAGHQADPES 203
 DB 124 CMTDKRVVNERGAIRAEHIREFWSAKSNINIERAFCELARAILDKTSGRESAENQ 183
 QY 204 FQIRDYVESQKKR---SSCCS 221
 DB 184 ERVIDRRNQEKAPGYSKCA 204
 RESULT 13
 Q40203
 ID Q40203 PRELIMINARY; PRT; 202 AA.
 AC Q40203;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE RAB1C.
 GN RAB1C.
 OS Lotus japonicus.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.
 OX NCBI_TaxID=34305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ROOT NODULES;
 RX MEDLINE-97231679; PubMed-9076991;
 RA Borg S., Brandstrup B., Jensen T.J., Poulsen C.;
 RT "Identification of new protein species among 33 different small GTP-
 binding proteins encoded by cDNAs from Lotus japonicus, and expression
 of corresponding mRNAs in developing root nodules.";
 RL Plant J. 11:237-250(1997).
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 DR EMBL; 273932; CAA98160.1; -;
 DR HSSP; P05713; 3RAB.
 DR InterPro; IPR003579; GTPase_Rab.
 DR InterPro; IPR001806; Ras_trnsfrmng.
 DR InterPro; IPR02078; Sig54_interact.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTRNSFRMNG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRFAMS; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
 KW GTP-binding; Lipoprotein.
 SQ SEQUENCE 202 AA; 22633 MW; 91633EBFAD07662F CRC64;
 Query Match 44.1%; Score 507.5; DB 10; Length 202;
 Best Local Similarity 46.8%; Pred. No. 7.6e-39;
 Matches 95; Conservative 46; Mismatches 55; Indels 7; Gaps 3;
 QY 23 SPSYDLTGKVMLLGDTGVGKTCFLIOFKDGAFLSGTFTIATVGDIDFRNKVVTVGVVRVKLQ 82
 DB 2 NPEYDLFKLLIGDSCGVGKTCILFRSDDAFTS-TFISTIGIDFKIKTVELRGKIKLQIW 60
 QY 83 TWDTAGQERFRSVTHAYYRDAQAALLLLYDITNKSSFNIRAWLTIEHYAQRDQVVMILG 142
 DB 61 INDWTAQERFRFTTSYRGAMGIMLYDITNKSSFNIRAWLTIEHYAQRDQVVMILG 120
 QY 143 NKADSSERVISGDTGLAREGVPPLETSAKTMVLAFLAIAKE-LKYRAGHQADPES 198
 DB 121 NKCDLTENKVSYSYETAKAFADIEGIPFWETSNAKATNVEQAFMAAAEIKNRMASQPVNN 180
 QY 199 ADEPSFQIRDYVESQKKRSCCS 221

DB 181 ARPPTVQIRGQPVNOK--SGCCS 201
 RESULT 14
 Q9LZD4
 ID Q9LZD4 PRELIMINARY; PRT; 216 AA.
 AC Q9LZD4;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE GTP-binding protein-like (Putative GTP-binding protein).
 GN F12E4_300.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S.,
 RA Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,
 RA Onodera C.S.; Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G.,
 RA Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,
 RA Koesema E., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M.,
 RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
 RA Southwick A.; Tracy S.E., Shinozaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RT "Full length cDNA of gene F12E4_300 (GI:7378637).";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
 RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
 RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E.,
 RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,
 RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
 RA Shinozaki K.; Davis R.W., Ecker J.R., Theologis A.;
 RT "Full length cDNA of gene F12E4_300 (GI:7378637).";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 DR EMBL; AL162751; CAB83313.1; -;
 DR EMBL; AY056351; AAL07200.1; -;
 DR EMBL; AY035134; AAK059629.1; -;
 DR HSSP; P05713; 3RAB.
 DR InterPro; IPR003579; GTPase_Rab.
 DR InterPro; IPR001806; Ras_trnsfrmng.
 DR InterPro; IPR002078; Sig54_interact.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTRNSFRMNG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRFAMS; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
 KW GTP-binding; Lipoprotein.
 SQ SEQUENCE 216 AA; 24037 MW; 50C7D3AF4DC86F67 CRC64;
 Query Match 44.1%; Score 507; DB 10; Length 216;
 Best Local Similarity 47.5%; Pred. No. 9.4e-39;
 Matches 103; Conservative 46; Mismatches 56; Indels 12; Gaps 4;
 QY 15 APERSPSPSYDLTGKVMLLGDTGVGKTCFLIOFKDGAFLSGTFTIATVGDIDFRNKVVTV 74
 DB 11 APERSPSPSYDLTGKVMLLGDTGVGKTCFLIOFKDGAFLSGTFTIATVGDIDFRNKVVTV 74

```
Db 4 APARA---RSDYDLIKLLIGDSGVKSCLLLRFSDDTFTT-SFITTIGIDFKIRTVEL 59
QY 75 DGVVKLIQWDTAGQERFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTIHEYAQR 134
Db 60 DGKRIKLIQWDTAGQERFRFTTAYYRGAMGILLVVDVDTDESSEFNIRNMKNIEQHASD 119
QY 135 DVVIMLLGNKADM-SSERVIRSEDEGTIAREYGVFPLETSAKTGMNVELAFIAIKELKY 193
Db 120 NVNKKILVGNKADMDESKRAVPTAKQALADEYGIKFFETSAKTNNLVNENVFMSIAKDIQ 179
QY 194 R-----AGHQADPEFSQIRDYVESOKRRSCCSEF 223
Db 180 RLTEETDKAEPQGIKITKQDTAASSSTAESACCSIV 216

RESULT 15
Q8VWF9
ID Q8VWF9 PRELIMINARY; PRT; 216 AA.
AC Q8VWF9;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Ras-related protein RAB8-4 (Ras-related protein RAB8-2).
GN RAB8-4 OR RAB8-2.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRIGHT YELLOW 2;
RA Torimoto N., Shimada K., Ito K., Yamamoto K.;
RT "Characterization of Rab8 from tobacco BY-2 cell.";
RL Submitted (JAN-2002) to the EMBL/Genbank/DDBJ databases.
DR EMBL; AB079023; BAB84325.1; -.
DR EMBL; AB079021; BAB84323.1; -.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsmng.
DR InterPro; IPR002078; Sig54_interact.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
SQ SEQUENCE 216 AA; 23971 MW; D5E87FABD0D49E13 CRC64;

Query Match 44.1%; Score 507; DB 10; Length 216;
Best Local Similarity 46.9%; Pred No. 9.4e-39;
Matches 100; Conservative 45; Mismatches 56; Indels 12; Gaps 4;

QY 19 SPP--CSPSYDLTGKYMLLGDTGVGKTCFLIOFKDGAFISGTFIATVGIDFRNKVYVDG 76
Db 3 APPARADYDYLKLLIGDSGVKSCLLLRFSDDTFTT-SFITTIGIDFKIRTIELDG 61
QY 77 VRVKLIQWDTAGQERFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTIHEYAQRDV 136
Db 62 KRILQIWDTAGQERFTTAYYRGAMGILLVVDVDTDESSEFNIRNMKNIEQHASDNV 121
QY 137 VIMLLGNKADM-SSERVIRSEDEGTIAREYGVFPLETSAKTGMNVELAFIAIKELKYRA 195
Db 122 NKILVGNKADMDESKRAVPTAKQALADEYGIKFFETSAKTNNLVNENVFMSIAKDIQRL 181
QY 196 GHQADPEFSQIRDYVESOK-----KRSSCC 220
Db 182 SEDSKTEPAIRINQSDQAGTSGQAQKSSCC 214
```

Search completed: November 17, 2002, 16:41:51
Job time : 82 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2002, 16:36:38 ; Search time 43 Seconds
(without alignments)
498.558 Million cell updates/sec

Title: US-09-817-199A-2

Perfect score: 1150

Sequence: 1 MTCTPGAVATRDGEAPERSP.....FQIRDYVESOKRRSSCCSFM 223

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	742	64.5	190	2 JC2528	GTP-binding protei
2	533	46.3	196	2 T15123	hypothetical prote
3	517.5	45.0	203	2 B34716	GTP-binding protei
4	516	44.9	206	2 I78851	GTP-binding protei
5	513.5	44.7	207	2 B49647	GTP-binding protei
6	513.5	44.7	207	2 B36364	GTP-binding protei
7	512.5	44.6	215	2 T14565	GTP-binding protei
8	507	44.1	216	2 T48378	GTP-binding protei
9	507	44.1	258	2 B86153	ARA-5 [imported] -
10	506	44.0	216	2 S33900	GTP-binding protei
11	505.5	44.0	201	2 T28971	hypothetical protei
12	500	43.5	202	2 S38740	GTP-binding protei
13	500	43.5	216	2 T45901	GTPase ATRAB8 - Ar
14	498	43.3	208	2 A34716	GTP-binding protei
15	497.5	43.3	202	2 S41430	GTP-binding protei
16	497.5	43.3	209	2 B38625	GTP-binding protei
17	497.5	43.3	215	2 S57478	GTP-binding protei
18	495	43.0	203	2 B38202	GTP-binding protei
19	495	43.0	216	2 S57471	GTP-binding protei
20	495	43.0	216	2 J50640	GTP-binding protei
21	494	43.0	200	2 D36364	GTP-binding protei
22	490.5	42.7	200	2 A38625	GTP-binding protei
23	490.5	42.7	215	2 S57462	GTP-binding protei
24	489	42.5	222	2 T14405	small GTP-binding
25	488	42.4	203	2 S34253	GTP-binding protei
26	487	42.3	200	2 S12790	GTP-binding protei
27	486	42.3	200	2 B42148	GTP-binding protei
28	486	42.3	202	2 S72515	GTP-binding protei
29	486	42.3	203	2 JC4105	GTP-binding protei

30	485	42.2	196	2 PS0279	GTP-binding protei
31	484.5	42.1	215	2 S57474	GTP-binding protei
32	483	42.0	221	2 H71444	GTP-binding protei
33	481.5	41.9	224	2 T33855	hypothetical prote
34	479.5	41.7	204	2 JC7589	Sec4p homolog - ye
35	479.5	41.7	208	2 A38202	GTP-binding protei
36	478.5	41.6	206	2 T14391	GTP-binding protei
37	478	41.6	201	2 S06147	GTP-binding protei
38	478	41.6	203	2 JC1247	GTP-binding protei
39	477.5	41.5	205	2 S38339	GTP-binding protei
40	476.5	41.4	205	2 T33781	hypothetical prote
41	476	41.4	217	2 S36365	GTP-binding protei
42	473.5	41.2	203	2 A49647	GTP-binding protei
43	468.5	40.7	201	2 S39565	GTP-binding protei
44	463.5	40.3	203	2 S51495	GTP-binding protei
45	459.5	40.0	203	2 S30096	GTP-binding protei

ALIGNMENTS

RESULT 1

JC2528

GTP-binding protein Rab26 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 02-Feb-2001

C:Accession: JC2528

R:Wagner, A.C.C.; Strowski, M.Z.; Goetze, B.; Williams, J.A.

Biochem. Biophys. Res. Commun. 207, 950-956, 1995

A:Title: Molecular cloning of a new member of the Rab protein family, Rab26, from rat

A:Reference number: JC2528; MUID:95169156; PMID:7864900

A:Accession: JC2528

A:Molecule type: mRNA

A:Residues: 1-190 <WAG>

A:Cross-references: GB:T08521

A:Note: The authors translated the codon GTG for residue 49 as Leu, GAA for residue 1

C:Genetics:

A:Gene: Rab26

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

C:Keywords: GTP binding; nucleotide binding; P-loop

F:4-11/Region: nucleotide-binding motif A (P-loop)

F:52-58/Region: GTP binding #status predicted

F:111-114/Region: GTP binding #status predicted

F:139-143/Region: GTP binding #status predicted

Query Match 64.5%; Score 742; DB 2; Length 190;
Best Local Similarity 72.9%; Pred. No. 3.4e-57;
Matches 137; Conservative 28; Mismatches 23; Indels 0; Gaps 0;

QY 33 MLLGDTGVGKTCFLQFDGAFSLGTFATVGDIFRNKVVTVDGVVRKLIQIWDTAGQERF 92

DB 1 MLVGSVGKGTCLLVRFKDGAFAGTFTSTVGDIFRNKVLVDVDMKVKVQIWDTAGQERF 60

QY 93 RSVTHAYRDAQALLLLDITNKSSFDNITRAWLTETHEYARQDVVIMLGNKADMSERV 152

DB 61 RSVTHAYRDAHALLLDITNKSSFDNITRAWLTETHEYARQDVVIMLGNKVDTSQERV 120

QY 153 IRSEGEITLARYGVPEFLETSKATGMNVELAFATAKELKYRAGHQADPEPSQIRDYVES 212

DB 121 VKREGEKLAKEYGLPFMETSAKSLGNVLDAFTATAKELKQRSTKAPSEPRFLHDYVKR 180

QY 213 QKKRSSCC 220

DB 181 EGRGVSCC 188

RESULT 2

T15123

hypothetical protein W01H2.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000

C:Accession: T15123

R:Minx, P.; Wohldmann, P.

submitted to the EMBL Data Library, April 1997
A:Description: The sequence of C. elegans cosmid W01H2.
A:Reference number: Z18296

A:Accession: T15123
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-196 <MIN>

A:Cross-references: EMBL:AF000192; NID:g1946982; PID:g1946985; PIDN:AAB52888.1; GSPDB:GN

A:Experimental source: strain Bristol N2; clone W01H2

C:Genetics:

A:Gene: CESP:W01H2.3

A:Map position: X

A:Introns: 41/3; 55/3; 95/3; 134/3; 160/2

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

F:3-119/Domain: translation elongation factor Tu homology <ETU>

Query Match 46.3%; Score 533; DB 2; Length 196;

Best Local Similarity 54.5%; Pred. No. 5.2e-39;

Matches 109; Conservative 34; Mismatches 41; Indels 16; Gaps 5;

Qy 31 KVMLLGDTGVGKTCFLIQKDGAFLSGTFTATVGVDFRNKVVTVDGVRVKLQIWDTAGOE 90

Db 4 KVMLLGDSCTGKTCCLIRKDKGAFNNFISVGVDFRNKLITMGDKVKLQIWDTAGOE 63

Qy 91 RFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTHEIHYAQRDVVIMLLGNKADMSSE 150

Db 64 RFRSVTTSYVRADALLLYDITANRASFCNRLWSQIKEYGKAVQVTLVGNKCDL--P 121

Qy 151 RVIRSDGETLAREYGVPLETSAGTKGMNVELAFIAIAK---ELKY---RAGHQADEPSF 204

Db 122 RAVPTDEGRRLAEAYQIPFMTSAGTKGFNVDRFLGLAERMLKLYGFPVGEMADTIS- 180

Qy 205 QIRDYVESQKR--SSCCSF 222

Db 181 ----VADTKKPEIARCTF 195

RESULT 3

B34716

C:Superfamily: Dictyostellium discoideum

C:Keywords: Dictyostellium discoideum

C:Date: 29-Jun-1990 #sequence_revision 29-Jun-1990 #text_change 02-Feb-2001

C:Accession: B34716; B61571

R:Saxe, S.A.; Kimmel, A.R.

Mol. Cell. Biol. 10, 2367-2378, 1990

A:Title: SAS1 and SAS2, GTP-binding protein genes in Dictyostellium discoideum with sequ

A:Reference number: A34716; MUID:90220623; PMID:2109188

A:Accession: B34716

A:Molecule type: DNA

A:Residues: 1-203 <SAX>

A:Cross-references: GB:M34457

R:Saxe, S.A.; Kimmel, A.R.

Dev. Genet. 9, 259-265, 1988

A:Title: Genes encoding novel GTP-binding proteins in Dictyostellium.

A:Reference number: A61571; MUID:89209367; PMID:3149563

A:Accession: B61571

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-203 <SA2>

C:Genetics:

A:Gene: SAS2

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

C:Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;

F:16-131/Domain: translation elongation factor Tu homology <ETU>

F:22-29/Region: nucleotide-binding motif A (P-loop)

F:128-131/Region: GTP-binding NKXD motif

F:158-160/Region: GTP-binding SAK/L motif

F:202,203/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match

Best Local Similarity 45.0%; Score 517.5; DB 2; Length 203;

Matches 97; Conservative 50; Mismatches 54; Indels 5; Gaps 3;

RESULT 5

B49647

C:Superfamily: GTP-binding protein rab8 - human

C:Keywords: Homo sapiens (man)

C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 19-Jan-2001

C:Accession: B49647; S36817

R:Zahraoui, A.; Joberty, G.; Arpin, M.; Fontaine, J.J.; Hellio, R.; Tavitian, A.; Lou

J. Cell Biol. 124, 101-115, 1994

A:Title: A small rab GTPase is distributed in cytoplasmic vesicles in non polarized c

Qy 15 APERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTFTATVGVDFRNKVVTV 74

Db 3 SPATNKPA--AYDFLVKLLIGDSGVGSKLLRSDGSF-TPSFATIGIDFKIRTIETL 59

Qy 75 DGVRVKLOIWDTAGOERFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTHEIHYAQR 134

Db 60 EGKRILQIWDTAGOERFTITAYRGAMGILLVDYDTEKSFGRINWIRNIIQHASD 119

Qy 135 DVVIMLLGNKADMSRVSIRSDGETLAREYGVPLETSAGTKGMNVELAFIAIAKELKYR 194

Db 120 SYNKMLIGNKCDMTEKKVVDSSRGKSLADEYGIKELETSAKNSVNVEEAFIGLANDIKKR 179

Qy 195 AGHQADEPSFQIRDYVESQKRSSCC 220

Db 180 MIDTPNDPDHTI--CITPNKKNTCC 203

RESULT 4

I78851

C:Superfamily: GTP-binding protein MEL - mouse

N:Alternate names: gene MEL protein

C:Species: Mus sp. (mouse)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 02-Feb-2001

C:Accession: I78851

R:Nimmo, E.R.; Sanders, P.G.; Padua, R.A.; Hughes, D.; Williamson, R.; Johnson, K.J.

Oncogene 6, 1347-1351, 1991

A:Title: The MEL gene: a new member of the RAB/YPT class of RAS-related genes.

A:Reference number: I58355; MUID:91360267; PMID:1886711

A:Accession: I78851

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-206 <RES>

A:Cross-references: GB:S53270; NID:g234747; PIDN:AAB19682.1; PID:g234748

C:Genetics:

A:Gene: MEL

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

C:Keywords: GTP binding; lipoprotein; nucleotide binding; P-loop; prenylated cysteine

F:9-124/Domain: translation elongation factor Tu homology <ETU>

F:15-22/Region: nucleotide-binding motif A (P-loop)

F:121-124/Region: GTP-binding NKXD motif

F:151-153/Region: GTP-binding SAK/L motif

F:203/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 44.9%; Score 516; DB 2; Length 206;

Best Local Similarity 49.5%; Pred. No. 1.7e-37;

Matches 101; Conservative 41; Mismatches 56; Indels 6; Gaps 3;

Qy 25 SYDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTFTATVGVDFRNKVVTVDVGVVKLQIW 84

Db 4 TYDYLFKLLIGDSGVGKTCVLFSEDAF-NSTFSTIGIDFKIRTIETLIDGKRILQIW 62

Qy 85 DTAGQERFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTHEIHYAQRDVVIMLLGNK 144

Db 63 DTAGQERFRFTITAYYRGAMGIMLVYDITNEKSFGRINWIRNIEHASADVEKMILGNK 122

Qy 145 AMSSRVSIRSDGETLAREYGVPLETSAGTKGMNVELAFIAIAKELKYRAGHQ----AD 200

Db 123 CDVNDKRQVSKERGEKALDYGIKFMETSAKANINVENAFFTLARDIKAKMDKNWKATAA 182

Qy 201 EPSFQIRDYVESQKRSSC-CSFM 223

Db 183 GSHGVKRVITVEQKRISFFRCULL 206

RESULT 5

B49647

C:Superfamily: GTP-binding protein rab8 - human

C:Species: Homo sapiens (man)

C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 19-Jan-2001

C:Accession: B49647; S36817

R:Zahraoui, A.; Joberty, G.; Arpin, M.; Fontaine, J.J.; Hellio, R.; Tavitian, A.; Lou

J. Cell Biol. 124, 101-115, 1994

A:Title: A small rab GTPase is distributed in cytoplasmic vesicles in non polarized c

A:Reference number: A49647; MUID:94124602; PMID:8294494

A:Accession: B49647

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-207 <ZAH>

A:Cross-references: EMBL:X56741; NID:g452317; PIDN:CAA0005.1; PID:g452318

R:Joberty, G.; Tavittian, A.; Zahraoui, A.

FEBS Lett. 330, 323-328, 1993

A:Title: Isoprenylation of Rab proteins possessing a C-terminal CaaX motif.

A:Reference number: S36817; MUID:93387463; PMID:8375503

A:Accession: S36817

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 175-186 <JOB>

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

C:Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;

F:9-124/Domain: translation elongation factor Tu homology <ETU>

F:15-22/Region: nucleotide-binding motif A (P-loop)

F:121-124/Region: GTP-binding NKXD motif

F:151-153/Region: GTP-binding SAK/L motif

F:204/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 44.7%; Score 513.5; DB 2; Length 207;

Best Local Similarity 50.0%; Pred. No. 2.8e-37;

Matches 99; Conservative 41; Mismatches 53; Indels 5; Gaps 2;

Qy 25 SYDLTGKVMLLGDTGVGKTCFLIQPKDGAFLSGTFTIATVGDIFRNKVVTVGVRVKLQIW 84

Db 4 TYDYLKLLIGDSGVGKTCVLFRESEDAF-NSTPISITIGIDFKIRTIELDGKRIKLQIW 62

Qy 85 DTAGQERPRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144

Db 63 DTAGQERPTTITAYYRGAMGIMLVYDITNEKSFNIRNIRNIEHASADVEKMILGNK 122

Qy 145 ADMSSERVIRSEDGTTLAREYGVPPLETSAKTMNVELAFIAIAKELKYRAGHQAD---- 200

Db 123 CDVNDKRVSKRGKLDYGIKFMETSAKININVENAFETLARDIKAKMDKKLEGNSP 182

Qy 201 EPSFOIRDYVESOKKRSS 218

Db 183 QGSNOGVKITPDQQRSS 200

Query Match 44.7%; Score 513.5; DB 2; Length 207;

Best Local Similarity 50.0%; Pred. No. 2.8e-37;

Matches 99; Conservative 41; Mismatches 53; Indels 5; Gaps 2;

Qy 25 SYDLTGKVMLLGDTGVGKTCFLIQPKDGAFLSGTFTIATVGDIFRNKVVTVGVRVKLQIW 84

Db 4 TYDYLKLLIGDSGVGKTCVLFRESEDAF-NSTPISITIGIDFKIRTIELDGKRIKLQIW 62

Qy 85 DTAGQERPRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144

Db 63 DTAGQERPTTITAYYRGAMGIMLVYDITNEKSFNIRNIRNIEHASADVEKMILGNK 122

Qy 145 ADMSSERVIRSEDGTTLAREYGVPPLETSAKTMNVELAFIAIAKELKYRAGHQAD---- 200

Db 123 CDVNDKRVSKRGKLDYGIKFMETSAKININVENAFETLARDIKAKMDKKLEGNSP 182

Qy 201 EPSFOIRDYVESOKKRSS 218

Db 183 QGSNOGVKITPDQQRSS 200

Query Match 44.7%; Score 513.5; DB 2; Length 207;

Best Local Similarity 50.0%; Pred. No. 2.8e-37;

Matches 99; Conservative 41; Mismatches 53; Indels 5; Gaps 2;

Qy 25 SYDLTGKVMLLGDTGVGKTCFLIQPKDGAFLSGTFTIATVGDIFRNKVVTVGVRVKLQIW 84

Db 4 TYDYLKLLIGDSGVGKTCVLFRESEDAF-NSTPISITIGIDFKIRTIELDGKRIKLQIW 62

Qy 85 DTAGQERPRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTEIHEYAQRDVVIMLLGNK 144

Db 63 DTAGQERPTTITAYYRGAMGIMLVYDITNEKSFNIRNIRNIEHASADVEKMILGNK 122

Qy 145 ADMSSERVIRSEDGTTLAREYGVPPLETSAKTMNVELAFIAIAKELKYRAGHQAD---- 200

Db 123 CDVNDKRVSKRGKLDYGIKFMETSAKININVENAFETLARDIKAKMDKKLEGNSP 182

Qy 201 EPSFOIRDYVESOKKRSS 218

Db 183 QGSNOGVKITPDQQRSS 200

Query Match 44.6%; Score 512.5; DB 2; Length 215;

Best Local Similarity 47.6%; Pred. No. 3.5e-37;

Matches 101; Conservative 47; Mismatches 53; Indels 11; Gaps 5;

Qy 19 SPP--CSPSYDLTGKVMLLGDTGVGKTCFLIQPKDGAFLSGTFTIATVGDIFRNKVVTVDG 76

Db 3 APPARADYDYLKLLIGDSGVGKSCLLLRFSGSEFTT-SFITTIGIDFKIRTIELDG 61

Qy 77 VRVQLQIWDTAGOERPRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTEIHEYAQRDV 136

Db 62 KRILQIWDTAGOERPTTITAYYRGAMGIMLVYDITDESSFNIRNIRNIEHASADNV 121

Qy 137 VIMLLGNKADM--SSERVIRSEDGTTLAREYGVPPLETSAKTMNVELAFIAIAKELKYRA 195

Db 122 NKILVGNKADMDESKRAVPTAKQALADYGIKFFETSAKTNLNVVEVFSIARDIKQRL 181

Qy 196 G----HQADPEPSFOIRDYVESOKK---RSSCC 220

Db 182 ADSDTROEAQPSITIKPADQSGNQAAKSACC 213

Query Match 44.6%; Score 512.5; DB 2; Length 215;

Best Local Similarity 47.6%; Pred. No. 3.5e-37;

Matches 101; Conservative 47; Mismatches 53; Indels 11; Gaps 5;

Qy 19 SPP--CSPSYDLTGKVMLLGDTGVGKTCFLIQPKDGAFLSGTFTIATVGDIFRNKVVTVDG 76

Db 3 APPARADYDYLKLLIGDSGVGKSCLLLRFSGSEFTT-SFITTIGIDFKIRTIELDG 61

Qy 77 VRVQLQIWDTAGOERPRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTEIHEYAQRDV 136

Db 62 KRILQIWDTAGOERPTTITAYYRGAMGIMLVYDITDESSFNIRNIRNIEHASADNV 121

Qy 137 VIMLLGNKADM--SSERVIRSEDGTTLAREYGVPPLETSAKTMNVELAFIAIAKELKYRA 195

Db 122 NKILVGNKADMDESKRAVPTAKQALADYGIKFFETSAKTNLNVVEVFSIARDIKQRL 181

Qy 196 G----HQADPEPSFOIRDYVESOKK---RSSCC 220

Db 182 ADSDTROEAQPSITIKPADQSGNQAAKSACC 213

Query Match 44.6%; Score 512.5; DB 2; Length 215;

Best Local Similarity 47.6%; Pred. No. 3.5e-37;

Matches 101; Conservative 47; Mismatches 53; Indels 11; Gaps 5;

Qy 19 SPP--CSPSYDLTGKVMLLGDTGVGKTCFLIQPKDGAFLSGTFTIATVGDIFRNKVVTVDG 76

Db 3 APPARADYDYLKLLIGDSGVGKSCLLLRFSGSEFTT-SFITTIGIDFKIRTIELDG 61

Qy 77 VRVQLQIWDTAGOERPRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTEIHEYAQRDV 136

Db 62 KRILQIWDTAGOERPTTITAYYRGAMGIMLVYDITDESSFNIRNIRNIEHASADNV 121

Qy 137 VIMLLGNKADM--SSERVIRSEDGTTLAREYGVPPLETSAKTMNVELAFIAIAKELKYRA 195

Db 122 NKILVGNKADMDESKRAVPTAKQALADYGIKFFETSAKTNLNVVEVFSIARDIKQRL 181

Qy 196 G----HQADPEPSFOIRDYVESOKK---RSSCC 220

Db 182 ADSDTROEAQPSITIKPADQSGNQAAKSACC 213

Query Match 44.6%; Score 512.5; DB 2; Length 215;

Best Local Similarity 47.6%; Pred. No. 3.5e-37;

Matches 101; Conservative 47; Mismatches 53; Indels 11; Gaps 5;

Qy 19 SPP--CSPSYDLTGKVMLLGDTGVGKTCFLIQPKDGAFLSGTFTIATVGDIFRNKVVTVDG 76

Db 3 APPARADYDYLKLLIGDSGVGKSCLLLRFSGSEFTT-SFITTIGIDFKIRTIELDG 61

Qy 77 VRVQLQIWDTAGOERPRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTEIHEYAQRDV 136

Db 62 KRILQIWDTAGOERPTTITAYYRGAMGIMLVYDITDESSFNIRNIRNIEHASADNV 121

Qy 137 VIMLLGNKADM--SSERVIRSEDGTTLAREYGVPPLETSAKTMNVELAFIAIAKELKYRA 195

Db 122 NKILVGNKADMDESKRAVPTAKQALADYGIKFFETSAKTNLNVVEVFSIARDIKQRL 181

Qy 196 G----HQADPEPSFOIRDYVESOKK---RSSCC 220

Db 182 ADSDTROEAQPSITIKPADQSGNQAAKSACC 213

Query Match 44.6%; Score 512.5; DB 2; Length 215;

Best Local Similarity 47.6%; Pred. No. 3.5e-37;

Matches 101; Conservative 47; Mismatches 53; Indels 11; Gaps 5;

Qy 19 SPP--CSPSYDLTGKVMLLGDTGVGKTCFLIQPKDGAFLSGTFTIATVGDIFRNKVVTVDG 76

Db 3 APPARADYDYLKLLIGDSGVGKSCLLLRFSGSEFTT-SFITTIGIDFKIRTIELDG 61

Qy 77 VRVQLQIWDTAGOERPRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLTEIHEYAQRDV 136

2

Db 6 YDMLFKLLIGDSGVGKTCILYRFSDDAF-NTTFSTIGIDFKIKTIELKGGKIKLIQWD 64
QY 86 TAGOERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTFIHEYAQRDVVIMLLGNKA 145
Db 65 TAGOERFRITTSYIRGANGIMLVYDITNAKSFDAKWLNRIDHSHASEDDVYKMITLGNKC 124
QY 146 DMSSESRVIRSEGETLAREYGVFPFLETSAKTMNVELAFALAKELKYRAGHQADPEPSFQ 205
Db 125 DMSDRVRVSRGEKIAQDHGHSFSETSAKLNHVHDTAFYDLAEAILAKMPDSTDQSRD 184
QY 206 INDYVESQKKRSS--CC 220
Db 185 TVNPVQPORQSSGGCC 201

RESULT 12
S38740
GTP-binding protein - rice
N:Alternate names: ras-related small GTP-binding protein
C:Species: Oryza sativa (rice)
C:Date: 27-May-1994 #sequence_revision 27-Feb-1997 #text_change 02-Feb-2001
C:Accession: S38740
R:Kidou, S.; Anai, T.; Umeda, M.; Aotsuka, S.; Tsuge, T.; Kato, A.; Uchimiya, H.
FEBS Lett. 332, 282-286, 1993
A:Title: Molecular structure of ras-related small GTP-binding protein genes of rice plant
A:Reference number: S38740; MUID:94009718; PMID:8405471
A:Accession: S38740
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-202 <KID>
A:Cross-references: GB:S66160; NID:g432606; PIDN:AAB28535.1; PID:g432607
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: GTP binding; nucleotide binding; P-loop
F:9-124/Domain: translation elongation factor Tu homology <ETU>
F:15-22/Region: nucleotide-binding motif A (P-loop)
F:121-124/Region: GTP-binding NRXD motif
F:151-153/Region: GTP-binding SAK/L motif
F:21,22,40,121,122,124,151/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 43.5%; Score 500; DB 2; Length 202;
Best Local Similarity 46.3%; Pred. NO. 4e-36;
Matches 94; Conservative 49; Mismatches 54; Indels 6; Gaps 3;
QY 23 SPSYDLTGKVMLLGDTGVGKTCFLIQKDFLGSCTFIATVGTIDPRNKVTVVDGVVRVKLQ 82
Db 2 NPEYDYLKLLIGDSGVGKSCLLLRFDSDYLE-SYSTIGVDFKIRVEODGKTIKLO 60
QY 83 IWDTAGOERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTFIHEYAQRDVVIMLLG 142
Db 61 IWDTAGOERFRITTSYIRGANGIMLVYDVTQDESFNVVKQWLNEIDRYASENVKLLVG 120
QY 143 NKADMSSESRVIRSEGETLAREYGVFPFLETSAKTMNVELAFALAKELKYRAGHQ---- 198
Db 121 NKCDLAENRVSVYEAGKALADIGIPFLETSAKDATNVEKAPMTWAGEIKNRMASQGRTN 180
QY 199 ADEPSFQIRDYVESQKKRSSCCS 221
Db 181 ASKPA-TVOMPRQPVQAQSSCCS 202

RESULT 13
T45901
GTPase ATRAB8 - Arabidopsis thaliana
N:Alternate names: protein F4P12.310
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 02-Sep-2000
C:Accession: T45901
R:Blocker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat, M.
submitted to the Protein Sequence database, January 2000
A:Reference number: 233016
A:Accession: T45901
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-216 <BLO>
A:Cross-references: EMBL:AL132966
A:Experimental source: cultivar Columbia; BAC clone F4P12
C:Genetics:
A:Map position: 3
A:Introns: 25/1; 49/2; 83/1; 105/1; 137/3; 158/3; 189/3
A:Note: F4P12.310
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
F:16-131/Domain: translation elongation factor Tu homology <ETU>

Query Match 43.5%; Score 500; DB 2; Length 216;
Best Local Similarity 46.5%; Pred. NO. 4.3e-36;
Matches 99; Conservative 45; Mismatches 57; Indels 12; Gaps 5;
QY 19 SPP--CSPSYDLTGKVMLLGDTGVGKTCFLIQKDFLGSCTFIATVGTIDPRNKVTVVDG 76
Db 3 APPARADYDYLKLLIGDSGVGKSCLLLRFDSDGFTT-SFTTIGIDFKIRTELDG 61
QY 77 VRVKLIQIWDTAGOERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTFIHEYAQRDV 136
Db 62 KRKLIQIWDTAGOERFRITTTAYRGAMGILLVYDVTDESSFNIRNIRNIEQHASDSV 121
QY 137 VTMLGNKADM--SSERVIRSEGETLAREYGVFPFLETSAKTMNVELAFALAKELKYRA 195
Db 122 NKILVGNKADMDESKRAVPKSKQALADEYGMKFFETSAKTNLVVEVFFSIKDKIKRL 181
QY 196 GH---QADEPSFQIRDY-----VESQKKRSSCC 220
Db 182 ADTDARAPQTIKINOSDQAGTSQATQKSACC 214

RESULT 14
A34716
GTP-binding protein SAS1 - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 29-Jun-1990 #sequence_revision 29-Jun-1990 #text_change 02-Feb-2001
C:Accession: A34716; A61571
R:Saxe, S.A.; Kimmel, A.R.
Mol. Cell. Biol. 10, 2367-2378, 1990
A:Title: SAS1 and SAS2, GTP-binding protein genes in Dictyostelium discoideum with se
A:Reference number: A34716; MUID:90220623; PMID:2109188
A:Accession: A34716
A:Molecule type: mRNA
A:Residues: 1-208 <SAX>
A:Cross-references: GB:M34456
R:Saxe, S.A.; Kimmel, A.R.
Dev. Genet. 9, 259-265, 1988
A:Title: Genes encoding novel GTP-binding proteins in Dictyostelium.
A:Reference number: A61571; MUID:89209367; PMID:3149563
A:Accession: A61571
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-208 <SA2>
C:Genetics:
A:Gene: SAS1
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop
F:16-131/Domain: translation elongation factor Tu homology <ETU>
F:22-29/Region: nucleotide-binding motif A (P-loop)
F:128-131/Region: GTP-binding NRXD motif
F:158-160/Region: GTP-binding SAK/L motif
F:207,208/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 43.3%; Score 498; DB 2; Length 208;
Best Local Similarity 46.6%; Pred. NO. 6.2e-36;
Matches 95; Conservative 49; Mismatches 50; Indels 10; Gaps 3;
QY 23 SPSYDLTGKVMLLGDTGVGKTCFLIQKDFLGSCTFIATVGTIDPRNKVTVVDGVVRVKLQ 82
Db 9 SAAYDYLIKLLIGDSGVGKSCLLLRFDSDSF-TPSFITTTIGIDFKIRTELEGGRIKLO 67
QY 83 IWDTAGOERFRSVTHAYYRDAQALLLLYDITNKSSFDNIRAWLTFIHEYAQRDVVIMLLG 142

```
|||||
Db 68 IWDTAGOERFTTTAYIRGAMGILLVYDVTDEKSGFNIRNWRNIEQHTSDSVNKNMLIG 127
QY 143 NKADMSERVIRSEDGETLAREYGVFPFLETSAKTGMNVELAFLAIKELKYRAGHQADEP 202
Db 128 NKCDMAEKVVDSRSGSLADEYGIKFLTSAKNSINVEAFISLAKDIKKR---MIDTP 184
QY 203 SFQIR-----DYVESQKRSSCC 220
Db 185 NEQPQVVQPGTNLCANNKKKACC 208

RESULT 15
S41430
GTP-binding protein, ras-like (clone vfa-ypt1) - fava bean
N:Alternate names: guanine nucleotide regulatory protein
C:Species: Vicia faba (fava bean)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 02-Feb-2001
C:Accession: S41430
R:Saalbach, G.; Thielmann, J.
submitted to the EMBL Data Library, January 1994
A:Description: Sequences of cDNA clones from cotyledons of Vicia faba encoding ypt/rab-x
A:Reference number: S41430
A:Accession: S41430
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-202 <SAA>
A:Cross-references: EMBL:295590; NID:g452358; PIDN:CAA82707.1; PID:g452359
A:Experimental source: clone vfa-ypt1
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop
F:9-124/Domain: translation elongation factor Tu homology <ETU>
F:15-22/Region: nucleotide-binding motif A (P-loop)
F:121-124/Region: GTP-binding NKXD motif
F:151-153/Region: GTP-binding SAK/L motif
F:21,22,40,121,122,124,151/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 43.3%; Score 497.5; DB 2; Length 202;
Best Local Similarity 45.8%; Pred. No. 6.6e-36;
Matches 93; Conservative 46; Mismatches 57; Indels 7; Gaps 3;

QY 23 SPYSYDLTGKVMILGDTGVGTCFLIOFKDGAFLSGTFIATVIGDFRNKVVTVGVRVKLQ 82
Db 2 NPEYDYLKLLIGDSGVGKSCLLRFADDSYID-SYISTIGVDFKIRTVQDQGTIKLQ 60
QY 83 IWDTAGOERFRSVTHAYRDAQALLLYDITNKSSFNIRAWLTEIHEYAQRDVIWMLG 142
Db 61 IWDTAGOERFTTSSYIRGAHGLIIVDVTDEESFNNKQWLSEIDRYASDNVKNLLVG 120
QY 143 NKADMSERVIRSEDGETLAREYGVFPFLETSAKTGMNVELAFLAIKELKYRAGHQ---- 198
Db 121 NKCDLTENRAVPYETAKAFADAEIGIPMETSAKDSINVEQAFMAMASSIKERMASQPTNN 180
QY 199 ADEPSPQIRDYVESQKRSSCCS 221
Db 181 ARPPTVQIRGQPVGQK--SGCCS 201
```

Search completed: November 17, 2002, 16:42:55
Job time : 44 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 17, 2002, 16:42:58 ; Search time 52 seconds
(without alignments)

1584.639 Million cell updates/sec

Title: US-09-817-199A-2

Perfect score: 1150

Sequence: 1 MTGTPGAVATDGEAPERSP.....FQINDYVESQKRSSCSF 223

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 328367 seqs, 184756068 residues

Total number of hits satisfying chosen parameters: 656734

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2.1/USPTO_spool/US09817199/runat_13112002_134903_5200/app_query.fasta_1.391
-DB=PublishedApplications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09817199.ecgn.1.1.20.erunat.13112002.134903.5200
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_NA:*

1:	/cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2.6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3:	/cgn2.6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2.6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2.6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2.6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7:	/cgn2.6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2.6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2.6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10:	/cgn2.6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11:	/cgn2.6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12:	/cgn2.6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13:	/cgn2.6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14:	/cgn2.6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1150	100.0	2674	10	US-09-817-199A-1
2	1145	99.6	1116	10	US-09-794-257-13
3	1140	99.1	875	12	US-10-051-986-10
4	977	85.0	576	10	US-09-794-257-15

Alignment Scores:
Pred. No.: 4.11e-136
Score: 1150.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 10

Length: 2674
Matches: 223
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

ALIGNMENTS

RESULT 1

US-09-817-199A-1

; Sequence 1, Application US/09817199A

; Patent No. US20020142380A1

; GENERAL INFORMATION:

; APPLICANT: SHAO, Wei et al.

; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE

; FILE REFERENCE: CLO01187

; CURRENT APPLICATION NUMBER: US/09/817,199A

; CURRENT FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 2674

; TYPE: DNA

; ORGANISM: Human

US-09-817-199A-1

US-09-817-199A-2 (1-223) x US-09-817-199A-1 (1-2674)

```
QY 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
Db 42 ATGACGGGACGCCAGCGCGCTTGCACCGGGATGGGAGGCCCGCCGAGCGCTCCCG 101
QY 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
Db 102 CCCTGCAGTCCGAGCTACGACCTACGGGCAAGGTGATCTCTGGGAGACACAGCGCTC 161
QY 41 GlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
Db 162 GGCAAAACATGTTCTCTGATCAATTCAAAGACGGGGCTTCTGTCCGGAACCTTCATA 221
QY 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
Db 222 GCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTGTGATGGATGGCGTGAGAGTGA 281
QY 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
Db 282 CTGCAGATCTGGGACACCGCTGGGCAGAACGGTTCGGAAGCGTCACCCATGCTTATTAC 341
QY 101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerPheAspAsn 120
Db 342 AGAGATGCTCAGGCGCTTGTCTGTGTATGACATCACCACAAATCTCTTCGACAAAC 401
QY 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
Db 402 ATCAGGCGCTGGCTCACTGAGATTCATGATGATGCCAGAGGACGTGTGTATCATGCTG 461
QY 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
Db 462 CTAGGCAACAAGCGGATATGAGCAGCAAGAGTATCGCTTCGGAAGACGAGAGACC 521
QY 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
Db 522 TTGGCAGGAGGTACCGGTGTCTCTCTGGAGCAGCGCCCAAGACTGGCATGAATGTG 581
QY 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
Db 582 GACTTAGCCCTTCTGCCATCGCCCAAGAACTGAATATCCGGCGCGGCATCAGCGCGAT 641
QY 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
Db 642 GAGCCCGAGCTTCAGATCCGAGACTATGATAGTCCAGAAAGCGCTCCAGAAAGCGCTC 701
QY 221 SerPheMet 223
Db 702 TCCTTCATG 710
```

RESULT 2

```
US-09-794-257-13
; Sequence 13, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1e1
; FILE OF INVENTION: Human G-Proteins
; FILE REFERENCE: 35800/209285
; CURRENT APPLICATION NUMBER: US/09/794, 257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (124)...(699)
US-09-794-257-13
```

Alignment Scores:

```
Pred. No.: 4,92e-136 Length: 1116
Score: 1145.00 Matches: 222
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.57% Indels: 0
DB: 10 Gaps: 0
```

US-09-817-199A-2 (1-223) x US-09-794-257-13 (1-1116)

```
QY 2 ThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro 21
Db 31 ACGGGACGCCAGCGCGCTTGCACCGGGATGGGAGGCCCGCCGAGCGCTCCCGCGCC 90
QY 22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41
Db 91 TCCAGTCCGAGCTACGACCTCAGGGCAAGGTGATGCTTCTGGGAGACACAGCGCTCGC 150
QY 42 LysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAla 61
Db 151 AAAACATGTTCTCTGATCAATTCAAAGACGGGGCTTCTCTCGGAACCTTCATAGCC 210
QY 62 ThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeu 81
Db 211 ACCGTCGGCATAGACTTCAGGAACAAGGTGGTGTGATGGCGTGAGTGAAGCTG 270
QY 82 GlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArg 101
Db 271 CAGATCTGGGACACCGCTGGGAGGACCGTTCGAGAGCTCACCCATGCTTATACAGA 330
QY 102 AspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIle 121
Db 331 GATGCTCAGCGCTTGTCTCTGTATGACATCACCACAAATCTCTTTCGACAAACATC 390
QY 122 ArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeu 141
Db 391 AGGGCGTGGCTCACTGAGATTCATGATGATGCCAGAGGAGCGTGGTGTATGCTGCTA 450
QY 142 GlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeu 161
Db 451 GGCAACAAGCGGATATGAGCAGCGAAGAGTATCCGTTCCGAGACGAGAGACCTTG 510
QY 162 AlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGlu 181
Db 511 GCGAGGAGTACGGTGTCTCTCTGGAGCAGCGCCCAAGACTGGCATGAATGTGGAG 570
QY 182 LeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGlu 201
Db 571 TTAGCCCTTCTGGCCATCGCCCAAGGAACCTGAATATCCGGCGCGGCATCAGCGGATG 630
QY 202 ProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCysSer 221
Db 631 CCCAGCTTCAGATCCGAGACTATGTAGAGTCCCAAGAAAGCGCTCCAGCTGCTGCTCC 690
QY 222 PheMet 223
Db 691 TTCATG 696
```

RESULT 3

```
US-10-051-986-10
; Sequence 10, Application US/10051986
; Patent No. US20020146770A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; Tang, Y. Tom
; Lal, Preeti
; Guegler, Karl J.
; Corley, Neil C.
; Patterson, Chandra
; Batra, Sajeev
; Baugho, Mariah R.
; TITLE OF INVENTION: RAS PROTEINS
```

NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
CURRENT APPLICATION NUMBER: US/10/051,986
FILING DATE: 15-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/766,551
FILING DATE: DECEMBER 12, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0168-1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 875 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: UCMCL5T01
CLONE: 1528559
SEQUENCE DESCRIPTION: SEQ ID NO: 10 :
US-10-051-986-10

Alignment Scores:
Pred. No.: 1.48e-135 Length: 875
Score: 1140.00 Matches: 221
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.13% Indels: 0
DB: 12 Gaps: 0

US-09-817-199A-2 (1-223) x US-10-051-986-10 (1-875)

QY 3 GlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProCys 22
|||||
DB 3 GGCAGCCAGGGCGGTGGCCACCGGGATGCGAGGCCCGCGCGCTCCCGCCCTGC 62
QY 23 SerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLys 42
|||||
DB 63 AGTCGCGAGTACGACCTCACGGCAAGTGATGCTTCGGGAGACACAGCGCTCGCAAA 122
QY 43 ThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThr 62
|||||
DB 123 ACATGTTTCCTGATCCAAATTCAGAGCGGGCTTCCTGCGGAACCTTCATAGCCACC 182
QY 63 ValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGln 82
|||||
DB 183 GTCGCGCATAGACTTCAGGAACAAGTGGTGACTGTGGATGGCGTGAGGTGAAGCTGCAG 242
QY 83 IleTrpAspThrAlaGlyClnGluArgPheArgSerValThrHisAlaTyrTyrArgAsp 102
|||||
DB 243 ATCTGGGACACCGCTGGGAGGACGGTTCGGAAGCTCACCCATGCTATTACAGAGAT 302
QY 103 AlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArg 122
|||||
DB 303 GCTCAGGCGCTGCTCTGTGTATGACATCATCAACAACAAATCTCTTCGACAAACATCAGG 362

QY 123 AlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValIleMetLeuLeuGly 142
|||||
DB 363 GCCTGGCTACATGAGATTCATGATGCCAGAGGGAGCTGGTGATCATGCTGCTAGGC 422
QY 143 AsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAla 162
|||||
DB 423 AACAAAGGCGGATATGAGCAGCGAAAGATGATCCGTTCCGAAAGACGAGAGACCTTGGCC 482
QY 163 ArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeu 182
|||||
DB 483 AGGAGTAGCGGTGTTCCCTTCTGGAGACCGCCCAAGACTGGCATGAATGTGGAGTTA 542
QY 183 AlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGluPro 202
|||||
DB 543 GCCTTCTGGCCATCGCCAAAGCACTGAATACCGGGCCGGCATCAGCGGATGAGCCC 602
QY 203 SerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysSerPhe 222
|||||
DB 603 AGCTTCCAGATCCGAGACTATGTAGAGTCCCGAGAGAGCGCTCCAGCTGCTGCTCTTC 662
QY 223 Met 223
DB 663 ATG 665

RESULT 4

US-09-794-257-15
; Sequence 15, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1e1
; TITLE OF INVENTION: Human G-Proteins
; FILE REFERENCE: 35800/209285
; CURRENT APPLICATION NUMBER: US/09/794,257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 576
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-794-257-15

Alignment Scores:
Pred. No.: 3.89e-115 Length: 576
Score: 977.00 Matches: 191
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 84.96% Indels: 0
DB: 10 Gaps: 0

US-09-817-199A-2 (1-223) x US-09-794-257-15 (1-576)

QY 33 MetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGly 52
|||||
DB 1 ATGCTCTCTGGGAGACACAGCGCTCGCAAAACATGTTTCTCTATCCANTTCAAGACGGG 60
QY 53 AlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValVal 72
|||||
DB 61 GCCTTCTGTCGGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTG 120
QY 73 ThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPhe 92
|||||
DB 121 ACTGTGGATGGGTGAGAGTGAAGCTGCAGATCTGGACACCGCTGGGAGGACCGTTC 180
QY 93 ArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAspIle 112
|||||
DB 181 CGAAGCGTCACCATGCTTATTACAGAGATGCTCAGGCGCTGCTTCTGCTGATGACATC 240
QY 113 ThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAla 132

Db 241 ACCAACAATCTCTTTCACACATCAGGGCTGCTCCTCAGATTCATGAGTAGTCC 300
Qy 133 GlnArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgVal 152
Db 301 CAGAGGAGCTGGTATCATGCTGCTAGGCAACAAGCGGATAGAGCAGCGAAGAGTG 360
Qy 153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThr 172
Db 361 ATCCGTTCCGAAGACGAGAGACCTTGGCCAGGAGTAGCGGTGTTCCCTTCCTGGAGACC 420
Qy 173 SerAlaLysThrClyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys 192
Db 421 AGCGCCAAGACTGGCAATGTTAGGAGTTAGCCTTCTGGCCATCGCCAAGGAACGTGAAA 480
Qy 193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212
Db 481 TACCGGGCCGGGCAATCAGCGGATGAGCCAGCTTCAGATCCGAGACTATGTAGAGTCC 540
Qy 213 GlnLysLysArgSerSerCysSerPheMet 223
Db 541 CAGAAGAAGCGCTCCAGCTGCTGCTCCTTCATG 573

RESULT 5

US-09-867-550-1811
; Sequence 1811, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1811
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: Wherein n is one of a or t or c or g
US-09-867-550-1811

Alignment Scores:
Pred. No.: 11e-82 Length: 447
Score: 720.00 Matches: 138
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 62.61% Indels: 0
DB: 10 Gaps: 0

US-09-817-199a-2 (1-223) x US-09-867-550-1811 (1-447)

Qy 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
Db 34 ATCAGGGCAGCCAGCGCGCTGTCACCCGGATGGCGAGGCCCGCCGAGCGCTCCCG 93
Qy 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
Db 94 CCCTGCAGTCCGAGCTACGACCTCACGGGCAAGGTATGCTTCCTGGAGACACAGCGCTC 153
Qy 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
Db 154 GGCAMACATGTTCTCTGATCCCAATTCAMAGACGGGGCCTTCCTGTCCGGACCTTCATA 213

Qy 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
Db 214 GCCACCGTCGGATAGACTTCAGGAACAAGTGGTGACTGTGGATGGCGTGAGAGTGAAG 273
Qy 81 LeuGlnIleTyrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
Db 274 CTGCAAGATCTGGACACCCCTGGCAGGAACGGTTCGGAAGCGTCACCCATGCTTATTAC 333
Qy 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
Db 334 AGAGATGCTCAGSCCTTGCTTCTGCTGATGACATCACCAACAATCTCTTCGACAAC 393
Qy 121 IleArgAlaTyrLeuThrGluIleHisGluTyrAlaGlnArgAspValValIle 138
Db 394 ATCAGGCGCTGCTCCTACTGAGATTCATGATGATGATGATGATGATGATGATGATGATG 447

RESULT 6

US-09-834-975-879
; Sequence 879, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF HUMAN CANCERS
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 879
; LENGTH: 2497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2497)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-879

Alignment Scores:
Pred. No.: 3.94e-55 Length: 2497
Score: 511.50 Matches: 106
Percent Similarity: 62.82% Conservative: 41
Best Local Similarity: 45.30% Mismatches: 68
Query Match: 44.48% Indels: 19
DB: 10 Gaps: 4

US-09-817-199a-2 (1-223) x US-09-834-975-879 (1-2497)

Qy 3 GlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro 21
Db 8 GGCACGCC-----CCTCGCCGCGCGCCCTCCCGCCCTCTCTCCACCGCCT 55
Qy 22 -----CysSerProSerTyrAspLeuThr 29
Db 56 CCTCTGGCTCCCGGTGTCAGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 115
Qy 30 GlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPhe 49
Db 116 TTCAACTCCTGCTGATCGCGACTCGGGGTAGGCAAGACCTCCCTCTGTTCCGCTTC 175
Qy 50 LysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsn 69
Db 176 TCAGAGGAGCGCTC---AACACCACTTCATCTCCACCATCGAATGATTTTAAAT 232
Qy 70 LysValValThrValAspGlyValArgValLysLeuGlnIleTyrPaspThrAlaGlyGln 89

```
Db 233 AGAAGCATAGAACTAGATGGAAGAAAATAAGCTTCAGATATGGACACACGCGGTGCG 292
Qy 90 GluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeu 109
Db 293 GAAAGATTCGCGACAACTCAGACACGCTACTACAGAGAGCCATGGCATATGCTGGTC 352
Qy 110 TyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHis 129
Db 353 TATGACATCACAAATGAAAAATCCTTTGACAATATTAATAAATTTGCATCAGAAACATTGAA 412
Qy 130 GluTyrAlaGlnArgAspValIleMetLeuLeuGlyAsnLysAlaAspMetSerSer 149
Db 413 GAGCATGCTCTTCGATGTCGAAAGAAATGATCCTGGGTGGAACAAATGTGATATGAATGAC 472
Qy 150 GluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPhe 169
Db 473 AAAAGACAAAGTCAAAAGAAAGAGGAGAGAGCTAGCAATGACTATGGGATTAATTC 532
Qy 170 LeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLys 189
Db 533 TTGGAGACAAAGCGCAAAATCCAGTGCAAATGTAGAAGAGGCATTTTACACTTGCACGA 592
Qy 190 GluLeuLysTyrArgAlaGlyHisGlnAla---AspGluProSerPheGlnIleArgasp 208
Db 593 GATATAAATGACAAAACTCAACAGAAAAATGAATGACAGCAATTCACAGAGCAGGTGGA 652
Qy 209 TyrValGluSerGlnLysLysArgSerSerCysCysSerPhe 222
Db 653 CCAGTGAANAATACAGAAACCGCATCAAGAAGACCAAGTTC 694

RESULT 7
US-09-834-975-885
; Sequence 885, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF HUMAN CANCERS
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 885
; LENGTH: 2497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2497)
US-09-834-975-885

Alignment Scores:
Pred. No.: 3,94e-55 Length: 2497
Score: 511.50 Matches: 106
Percent Similarity: 62.82% Conservative: 41
Best Local Similarity: 45.30% Mismatches: 68
Query Match: 44.48% Indels: 19
DB: 10 Gaps: 4

US-09-817-199A-2 (1-223) x US-09-834-975-885 (1-2497)

Qy 3 GlyThrProGlyAlaValAlaThrArgaspGlyGluAlaProGluArgSerProPro--- 21
Db 8 GGCACGCC-----CCTCGCGCGCGCGGCCCTCCCGCGCTCTCTCCACGCGCT 55
```

```
Qy 22 -----CysSerProSerTyrAspLeuThr 29
Db 56 CCTCTGGTCCCGCTCAGAGGCGCGAGAGAGATGGCGAAGACGTACGATTATCTC 115
Qy 30 GlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuIleGlnPhe 49
Db 116 TTCAAGCTCCTCTGATCGGCGACTCGGGGTAGCAAGACCTGCTCTGTTCCGCTTC 175
Qy 50 LysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsn 69
Db 176 TCAGAGGAGCGCTTC---AACACCACTTCATCTCCACCTCGGAATGATTTTAAAT 232
Qy 70 LysValValThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGln 89
Db 233 AGAAGCATAGAACTAGATGGAAGAAAATAAGCTTCAGATATGGACACACGCGGTGCG 292
Qy 90 GluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeu 109
Db 293 GAAAGATTCGCGACAACTCAGACACGCTACTACAGAGAGCCATGGCATATGCTGGTC 352
Qy 110 TyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHis 129
Db 353 TATGACATCACAAATGAAAAATCCTTTGACAATATTAATAAATTTGCATCAGAAACATTGAA 412
Qy 130 GluTyrAlaGlnArgAspValIleMetLeuLeuGlyAsnLysAlaAspMetSerSer 149
Db 413 GAGCATGCTCTTCGATGTCGAAAGAAATGATCCTGGGTGGAACAAATGTGATATGAATGAC 472
Qy 150 GluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPhe 169
Db 473 AAAAGACAAAGTCAAAAGAAAGAGGAGAGAGCTAGCAATTCGACTATGGGATTAATTC 532
Qy 170 LeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLys 189
Db 533 TTGGAGACAAAGCGCAAAATCCAGTGCAAATGTAGAAGAGGCATTTTACACTTGCACGA 592
Qy 190 GluLeuLysTyrArgAlaGlyHisGlnAla---AspGluProSerPheGlnIleArgasp 208
Db 593 GATATAAATGACAAAACTCAACAGAAAAATGAATGACAGCAATTCACAGAGCAGGTGGA 652
Qy 209 TyrValGluSerGlnLysLysArgSerSerCysCysSerPhe 222
Db 653 CCAGTGAANAATACAGAAACCGCATCAAGAAGACCAAGTTC 694

RESULT 8
US-09-834-975-894
; Sequence 894, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF HUMAN CANCERS
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 894
; LENGTH: 2497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2497)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-894
```

```

Alignment Scores:      3.94e-55      Length:      2497
Pred. No.:           511.50       Matches:     106
Score:               62.82%       Conservative: 41
Percent Similarity:  45.30%       Mismatches:   68
Best Local Similarity: 44.48%      Indels:       19
Query Match:         10          Gaps:         4
DB:

US-09-817-199A-2 (1-223) x US-09-834-975-894 (1-2497)

Qy 3 GlyThrProGlyAlaValAlaLathrArgAspGlyGluAlaProGluArgSerProPro--- 21
    |||||||
Db 8 GGCACGCC-----CCTCGCCGGCCCCCTCCCCTCTCTCCACCGCCT 55

Qy 22 -----CysSerProSerTyrAspLeuThr 29
    :|||:

Db 56 CCTCTGGCTCCCCGTTCAGAGGGCCGAGAGATGGCAAGCATGGATTATATCTC 115
    |||||||

Qy 30 GlyLysValMetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuLeuGlnPhe 49
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 116 TTCAGCTCTCTGATCGCGACTCGGGGTAGGACAGACCTGCTCTCTGTTCCGCTTC 175

Qy 50 LysAspGlyAlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAsn 69
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 176 TCAGAGGAGCGCTTC--AACACCACCTTCATCCACCATCGGAATTGATTTAAAT 232

Qy 70 LysValValThrValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGln 89
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 233 AGAACATAGAATAGTAGTAAGAAAATAAGCTTCAGATATGGACACAGCGGTTCAG 292

Qy 90 GluArgPheArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeu 109
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 293 GAAGATTTCGAACATACGACAGCGTACTACAGAGGAGCCATGGCATTTATGCTGGTC 352

Qy 110 TyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluLeuHis 129
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 533 TTGGAGACAAGCGCAAATCCAGTCGCAATGTAGAAGGCGATTTTTTACACTTGCACGA 592

Qy 190 GluLeuLysTyrArgAlaGlyHisGlnAla---AspGluProSerPheGlnIleArgAsp 208
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 593 GATATAATGACAAACTCAACAGAAAATGAATGACAGCAATTCACAGGAGCAGGTGGA 652

Qy 209 TyrValGluSerGlnLysLysArgSerSerCysCysSerPhe 222
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 653 CCAGTGAANAATACAGAAACCGATCAAGAGACCGCTTC 694

RESULT 9
US-09-834-975-896
; Sequence 896, Application US/09034975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975

```



```
Db 204 GACACAGCGGGTCAGAAAGATTCGGAACAATCACGACAGCGCTACTACAGAGGCCCATG 263
Qy 105 AlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArgAlaTrp 124
Db 264 GCATTATGCTGCTCTATGACATCAACAATGAAANAATCCTTTGACAATATTAANAATGG 323
Qy 125 LeuThrGluIleHisGluTyrAlaGlnArgAspValAlaIleMetLeuLeuGlyAsnLys 144
Db 324 ATCAGAAACATGGAAGAGCATGCCCTCTCCGATGTCGAAAGAATGATCTGGGTAAACAA 383
Qy 145 AlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAlaArgGlu 164
Db 384 TGTGATATCAATGACAAAGACAGTGTCTCCGATGTCGAAAGAATGATCTGGGTAAACAA 443
Qy 165 TyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAlaPhe 184
Db 444 TATGGGATTAATCTTGGAGACAGCGCAANAATCCAGTGCAANAATGTAGAAGGCAATTT 503
Qy 185 LeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAla---AspGluProSer 203
Db 504 TTTACACTTGCACGAGATATATGACAAAACCTCAACAGAAAATCAATGACAGCAATTC 563
Qy 204 PheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysSerPhe 222
Db 564 GCAGGAGCAGGTGGACCAAGTGAATAAACAAGAAAACCGATCAAGAAAGACCAAGTTTC 620

RESULT 14
US-09-938-842A-836
; Sequence 836, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 836
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-836

Alignment Scores:
Pred. No.: 1,58e-54 Length: 651
Score: 500.00 Matches: 99
Percent Similarity: 67.61% Conservative: 45
Best Local Similarity: 46.48% Mismatches: 57
Query Match: 43.48% Indels: 12
DB: 9 Gaps: 5

US-09-817-199A-2 (1-223) x US-09-938-842A-836 (1-651)
Qy 19 SerProTyrAspLeuThrGlyLysValMetLeuLeuGly 36
Db 7 GCTCTCTCTGCTGAGAGCTGCTGCTGATTACCTCAATTAACCTTCTGCTGATCGGA 66
Qy 37 AspThrGlyValGlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSer 56
Db 67 GACAGCGGTGGTGAAGAGTTGCCCTCTTACAGATTCTTCAGATGGCTGCTTACCACC 126
Qy 57 GlyThrPheIleAlaThrValGlyIleAspPheArgAsnLysValIleThrValAspGly 76
```

```
Db 127 ---ACCTTTTCATTACAACTATTGGGATTGATTGTTAAGATACGCACTATTGAGTCTGATGG 183
Qy 77 ValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThr 96
Db 184 AAGACAANTTAAGCTGCAAAATCTGGGATCTGCGGACAGAGCGGTTCGCCACAATCACA 243
Qy 97 HisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSer 116
Db 244 ACTGCGTACTACCGTGGAGCCATGGGATTGTTGCTGTATGATGCTGACTGATGAATCA 303
Qy 117 SerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspVal 136
Db 304 TCTTTCAACAACATCAGGAATTTGGATTCGTAACATTTAGCAGCATGCTTCTGATAGTGC 363
Qy 137 ValIleMetLeuLeuGlyAsnLysAlaAspMet---SerSerGluArgValIleArgSer 155
Db 364 AACAGATTCTAGTTGGGAACAACAGATATGGATGGAAGAAAGAAAGACTGTGCCAAAA 423
Qy 156 GluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLys 175
Db 424 TCTAAGGSCCAAGCTCTTGCAGATGAATATGAATGAAGTTTTCGAGACTAGTGCCAAG 483
Qy 176 ThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAla 195
Db 484 ACTAACTTAAAGCTTGAGGAAGTTTCTCTCTATTGCTTAAAGACATTAAGCAAGACTT 543
Qy 196 GlyHis-----GlnAlaAspGluProSerPheGlnIleArgAspTyr----- 209
Db 544 GCAGATACCGATGCAGAGCTGAGCGGCAACAACAAATCAACCAATCCGACCAAGGT 603
Qy 210 -----ValGluSerGlnLysLysArgSerSerCysCys 220
Db 604 GCGGGAACATCTCAACGACTCAGAAATCAGCATGCTTGC 642

RESULT 15
US-09-938-842A-832
; Sequence 832, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 832
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-832

Alignment Scores:
Pred. No.: 3,98e-54 Length: 609
Score: 496.50 Matches: 93
Percent Similarity: 68.97% Conservative: 47
Best Local Similarity: 45.81% Mismatches: 56
Query Match: 43.17% Indels: 7
DB: 9 Gaps: 3

US-09-817-199A-2 (1-223) x US-09-938-842A-832 (1-609)
Qy 23 SerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGlyLys 42
```

```

Db      4  AATCCTGAATATGACATATTGTTCAAGCTTTGCTCATTGGTGATTTCTGGTTGGAAG 63
QY      43  ThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAlaThr 62
Db      64  TCGTGTGCTTCTTAAGTTTGTGTATGATTCCTACTCTGGAT---AGCTACATCAGCAC 120
QY      63  ValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLysLeuGln 82
Db      121  ATTTGGTGTCTGACATTTAAATCCGACATCGAACAAGATCGAAAGACCATCAAACTCCAG 180
QY      83  IleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAsp 102
Db      181  ATTTGGGATACGGCAGCCAGGAACGTTTCAGAACGATTACTAGCAGTTACTACAGAGGA 240
QY      103  AlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArg 122
Db      241  GCTCATGGAAATTAATGTGACTTAATGATGTAACAGATCTAGAAAGCTTCAACAACGTCAG 300
QY      123  AlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeuGly 142
Db      301  CAATGGCTAAATGAAATCGACCGCTATGCAAGTGAGAATGTTAACAAGCTACTGTTGGG 360
QY      143  AsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAla 162
Db      361  AACAAAGTGTGATCTCATCATCACAAAGTTGTATCCACTGAGACAGCTAAGGCTTTGCT 420
QY      163  ArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeu 182
Db      421  GATGAACCTGGATCCCATCTCTTGGAAACAAGTGCTAAGAAATGCTACCAATGTCGAGAA 480
QY      183  AlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGln----- 198
Db      481  GCTTTCATGCCATGACTGCTGCAATCAAGACCAGATGCGAGCCAACTGCAGGAGGA 540
QY      199  AlaAspGluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSer 218
Db      541  TCTAAGCCACCAACGGTCCAGATCCGAGGACAACCTGTTTAACCAG-----CAATCAGGC 594
QY      219  CysCysSer 221
Db      595  TGCTGCTCC 603
```

Search completed: November 17, 2002, 18:06:57
Job time : 57 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 17, 2002, 16:40:13 ; Search time 254 seconds
(without alignments)
1977.148 Million cell updates/sec

Title: US-09-817-199A-2
Perfect score: 1150
Sequence: 1 MTGPPGAVATRDGEAPERSP.....FQIRDYVESQKRSCCSFM 223

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-O=/cgn2_1/USPFO_spool/US09817199/runat_13112002_134901_5153/app_query.fasta_1.391
-DB=N_Geneseq_101002 -OFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTENT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09817199 -RCGN_1_1_125 @runat_13112002_134901_5153 -NCFU=6 -ICPU=3
-NO_XLPXY -NO_MAP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FPGAPOP=6 -FPGAEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1150	100.0	1106	22	AAI59066	Human polynucleoti
2	1150	100.0	2323	22	AAI60852	Human polynucleoti
3	1145	99.6	1109	22	AAH75184	Nucleotide sequenc
4	1144	99.5	2623	22	AAH27036	cDNA encoding nove
5	1140	99.1	875	24	ABN83696	Human Ras protein
6	1133	98.5	1316	22	AAH27458	cDNA encoding nove
7	1034	89.9	843	20	AAH86720	DNA sequence encod
8	969	84.3	576	24	ABA02774	Human degranulatio
9	942	81.9	576	24	ABA02773	Mouse degranulatio
10	770	67.0	475	23	AAH81431	DNA encoding novel
11	731	63.6	1340	19	AAH65197	Human RAB protein,
12	697	60.6	2001	23	ABL29523	Drosophila melanog
13	535	46.5	964	22	AAH27040	cDNA encoding nove
14	516	44.9	911	21	AAH38429	Arabidopsis thalia
15	515.5	44.8	1129	21	AAH44482	Zea mays DNA fragm
16	515	44.8	1203	21	AAH43685	Zea mays DNA fragm
17	514.5	44.7	2411	24	ABQ54410	Human ovarian anti
18	513.5	44.7	1986	22	ABA09160	Human rab8 homolog
19	511.5	44.5	2497	22	AAH60878	Human cancer agent
20	511.5	44.5	2497	22	AAH60884	Human cancer agent
21	511.5	44.5	2497	22	AAH60893	Human cancer agent
22	511.5	44.5	2497	22	AAH60895	Human cancer agent
23	511.5	44.5	2497	23	ABV25781	Human prostate exp
24	511.5	44.5	2497	23	ABV30037	Human prostate exp
25	508.5	44.2	1540	23	ABL27707	Drosophila melanog
26	507	44.1	777	21	AAH42684	Arabidopsis thalia
27	507	44.1	881	24	ABN98761	Arabidopsis thalia
28	507	44.1	1023	21	AAH51491	Arabidopsis thalia
29	507	44.1	1025	21	AAH34347	Arabidopsis thalia
30	507	44.1	3077	21	AAH13912	Human cDNA sequenc
31	506.5	44.0	956	22	AAH96887	Nucleotide sequenc
32	506	44.0	674	23	AAH71453	DNA encoding novel
33	506	44.0	1537	21	AAH16196	Human prostate can
34	506	44.0	3533	22	AAH17889	Human cDNA sequenc
35	505.5	44.0	716	21	AAA40104	Human Rab10 cDNA.
36	505.5	44.0	861	21	AAA40108	Human Rab10 cDNA #
37	503	43.7	959	21	AAH33987	Arabidopsis thalia
38	502	43.7	1101	21	AAH34080	Arabidopsis thalia
39	501	43.6	866	22	AAH04301	Human cDNA clone (
40	501	43.6	1161	22	AAH75182	Nucleotide sequenc
41	498	43.3	888	21	AAA40109	Canine Rab10 cDNA.
42	497.5	43.3	932	21	AAH35200	Arabidopsis thalia
43	490	42.6	2247	23	ABL29661	Drosophila melanog
44	488	42.4	705	21	AAH49208	Arabidopsis thalia
45	483	42.0	666	21	AAH42764	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAI59066
ID AAI59066 standard; cDNA; 1106 BP.
XX
AC AAI59066;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1269.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.

XX WPI; 2001-442253/47.
DR P-PSDB; AAM41696.
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 4841; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 2323 BP; 542 A; 646 C; 613 G; 521 T; 1 other;

Alignment Scores:
Pred. No.: 6,74e-128 Length: 2323
Score: 1150.00 Matches: 223
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-09-817-199A-2 (1-223) x AAI60852 (1-2323)

Qy 1 MetThrGlyThrProGlyValAlaAlaThrArgAspGlyGluAlaProGluArgSerPro 20
Db 28 ATGACGGGACGCCAGGCGCGTTGCCACCGGGATGGCGAGGCCCGCGCTCCCGC 87
Qy 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
Db 88 CCTTCAGTCGCGAGCTACGACCTCACGGGCAAGGTGATGCTTCTGGGAGACACAGCGTC 147
Qy 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
Db 148 GGCRAAACATGTTTCTCTGATCCAAATTCAAAGACGGGGCTTCTCTCGGAACCTTCATA 207
Qy 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
Db 208 GCCACCGTCGGCATAGACTTTCAGGAACAAAGGTGGTGAATGGCGTGAGATGAAG 267
Qy 81 LeuGlnIleTyrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
Db 268 CTGCAGATCTGGGACACCGCTGGCGAGCAACGGTTCCCAAGCGTCACCCATGCTTATTAC 327
Qy 101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
Db 328 AGAGATGCTCAGGGCTTCTCTGCTGTATGACATCACCAACAATCTTCTTCGACAC 387
Qy 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
Db 388 ATCAGGGCTGGCTCCTCAGATTCATGATGATGCCAGAGGGAGCTGGTGATCATGCTG 447
Qy 141 LeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThr 160
Db 448 CTAGGCAACAGGGCGATAGCAGCGCAAGAGGTATCCCTCCGAGACGGAGAC 507
Qy 161 LeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnVal 180
Db 508 TTGGCCAGGGAGTACGGTGTTCCTCTCTGGAGACCGACCGCAAGACTGGCATGATG 567

Qy 181 GluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAsp 200
Db 568 GAGTTAGCCTTCTGGCCATCGCCAAAGCACTGAATACCGGGCGGCATCAGCGGAT 627
Qy 201 GluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
Db 628 GAGCCAGCTTCCAGACTCCGAGACTATGTAGTGTCCAGAGAGCGTCCAGCTGCTGC 687
Qy 221 SerPheMet 223
Db 688 TCCTTTCATG 696
RESULT 3
AAH75184
ID AAH75184 standard; DNA; 1109 BP.
XX AAH75184;
AC AAH75184;
XX
DT 13-NOV-2001 (first entry)
XX
DE Nucleotide sequence of human 32712 G-protein.
XX
KW Human; G-protein; 32705; 23224; 27423; 32700; 32712; lung disorder;
KW congenital anomaly; pulmonary congestion; oedema; haemorrhage;
KW adult respiratory distress syndrome; Goodpasture's syndrome;
KW chronic obstructive pulmonary disease; asthma; pulmonary hypertension;
KW liver disorder; hepatic injury; jaundice; cholestasis; viral hepatitis;
KW cirrhosis; Wilson's disease; autoimmune hepatitis; hepatic failure;
KW brain disorder; hypoxia; cerebral ischemia; intracranial haemorrhage;
KW acute meningitis; Parkinson's disease; Alzheimer's disease; glioma;
KW chronic bacterial meningencephalitis; multiple sclerosis;
KW amyoctic lateral sclerosis; stroke; Huntington's disease; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 124..699
FT FT /*tag= a
FT FT /product= "G-protein"
XX
PN WO200164887-A2.
XX
PD 07-SEP-2001.
XX
PF 27-FEB-2001; 2001WO-US06292.
XX
PR 29-FEB-2000; 2000US-0185606.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Meyers RA;
XX
DR WPI; 2001-550182/61.
XX P-PSDB; AAG67156.
PT Novel human small G-protein polypeptides and polynucleotides for
PT treating lung disorders, liver disorders and brain disorders -
XX
PS Claim 2; Fig 26; 151pp; English.
XX
CC The present sequence encodes a human G-protein. The specification
CC describes 32705, 23224, 27423, 32700 or 32712 small G-proteins. The
CC G-protein polypeptides and polynucleotides are useful as a target for
CC diagnosis and treatment of G-protein mediated or related disorders,
CC and for identifying agonists and antagonists for diagnosis and
CC treatment. They are useful for treating disorders of lung (e.g.
CC congenital anomalies, pulmonary congestion, oedema, adult respiratory
CC distress syndrome, haemorrhage, chronic obstructive pulmonary disease,
CC asthma, Goodpasture's syndrome and pulmonary hypertension), liver
CC (e.g. hepatic injury, jaundice, cholestasis, viral hepatitis, cirrhosis,
CC Wilson's disease, autoimmune hepatitis and hepatic failure), and
CC brain (e.g. hypoxia, cerebral ischemia, intracranial haemorrhage, acute

CC meningitis, Parkinson's disease, Alzheimer's disease, gliomas, chronic
CC bacterial meningoencephalitis, multiple sclerosis, amyotrophic lateral
CC sclerosis, stroke and Huntington's disease).
XX
SQ Sequence 1109 BP; 259 A; 309 C; 321 G; 220 T; 0 other;

Alignment Scores:
Pred. No.: 9,45e-128 Length: 1109
Score: 1145.00 Matches: 222
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.57% Indels: 0
DB: 22 Gaps: 0

US-09-817-199a-2 (1-223) x AAH75184 (1-1109)

QY 2 ThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro 21
Db 31 ACGGGACACGCCAGCCGCTTGCACCGGATGGCGAGGCCGCCCGAGCGCTCCCGGCC 90
QY 22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41
Db 91 TGCAGTCCGAGCTACGACCTCAGGGCAAGGTGATCTTCTGGAGACACAGCGCTCGC 150
QY 42 LysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIleAla 61
Db 151 AAAACATGTTCTGTATCAAAATTCAAAGACGGGGCTTCTGTCCGGAACCTTCATAGCC 210
QY 62 ThrValGlyLeuAspPheArgAsnLysValValThrValAspGlyValArgValLysLeu 81
Db 211 ACCGTGGCATGACATTCAGAACAGGTGGTGTGTGTGGTGGTGGTGGTGGTGGTGGT 270
QY 82 GlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaIleTyrArg 101
Db 271 CAGATCTGGACACCCCTGGCAGGAACGGTTCGGAAGCGTCACCCATGCTTATTACAGA 330
QY 102 AspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIle 121
Db 331 GATGCTCAGGCGCTTGTCTGTGTATGATCATCCACCAAAATCTTCTTCGACAAATC 390
QY 122 ArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeu 141
Db 391 AGGCGCTGCTCACTCAGATTCTATGATGATGCCAGAGGAGCGTGTGTATCATCTGCTA 450
QY 142 GlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeu 161
Db 451 GGCAACAAGCGGATATGACAGCAGGAAGAGTATCCGTCGAGACGAGAGACCTTG 510
QY 162 AlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGlu 181
Db 511 GCCAGGAGTACGGTGTCTCCCTTCCTGGAGACCGCCCAAGACTGGCATGATGGAG 570
QY 182 LeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGlu 201
Db 571 TTAGCTTCTCGCCATCGCCAGGAGACTGAATATCCGGCGGGCATCAGCGGATGAG 630
QY 202 ProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCysSer 221
Db 631 CCCAGCTTCCAGATCCGAGACTATGTAGATGCCAGAAGAGCGCTCCAGCTGCTGCTCC 690
QY 222 PheMet 223
Db 691 TTCATG 696

RESULT 4
AAS27036
ID AAS27036 standard; cDNA; 2623 BP.
XX
AC AAS27036;
XX
DT 07-NOV-2001 (first entry)
XX
DE cDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; rheumatoid arthritis; inflammatory condition;
KW immune system disorder; infection; hepatitis C; blood disorder;
KW organ transplant rejection; hyperproliferative disorder; Gaucher's disease;
KW sickle cell anaemia; Alzheimer's disease; Parkinson's disease;
KW neurodegenerative disorder; Down syndrome; ischaemia; renal disorder;
KW chromosomal abnormality; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
XX acquired immune deficiency syndrome.
OS Homo sapiens.
XX
PN WO200154733-A1.
XX
XX 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01312.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.

QY	203	SerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCysSerPhe	222
Db	603	ACGCTCCAGATCCGAGACTATGTAGAGTCCCAAGAGCGCTCCAGCTGCTGCTCCTC	662
QY	223	Met	223
Db	663	ATC	665
RESULT	6		
AAS27458			
ID	AAS27458	standard; cDNA; 1316 BP.	
AC	AAS27458;		
XX			
XX			
DT	07-NOV-2001	(first entry)	
DE		cDNA encoding novel signal transduction pathway protein, Seq ID 493.	
XX			
KW		Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;	
KW		antiflammatory; anti-HIV; antibacterial; antinflammatory; cancer;	
KW		immune system disorder; rheumatoid arthritis; inflammatory condition;	
KW		organ transplant rejection; infection; hepatitis C; blood disorder;	
KW		sickle cell anaemia; hyperproliferative disorder; Parkinson's disease;	
KW		neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;	
KW		chromosomal abnormality; Down syndrome; ischaemia; renal disorder;	
KW		cardiovascular; respiratory; wound healing; endocrine; Addison's disease;	
KW		reproductive system; gastrointestinal; liver disorder; AIDS; ss;	
KW		acquired immune deficiency syndrome.	
XX			
OS		Homo sapiens.	
XX			
PN	WO200154733-A1.		
XX			
PD	02-AUG-2001.		
XX			
PF	17-JAN-2001; 2001WO-US01312.		
XX			
PR	31-JAN-2000; 2000US-0179065.		
PR	04-FEB-2000; 2000US-0180628.		
PR	24-FEB-2000; 2000US-0184664.		
PR	02-MAR-2000; 2000US-0186350.		
PR	16-MAR-2000; 2000US-0189874.		
PR	17-MAR-2000; 2000US-0190076.		
PR	18-APR-2000; 2000US-0198123.		
PR	19-MAY-2000; 2000US-0205515.		
PR	07-JUN-2000; 2000US-0209467.		
PR	28-JUN-2000; 2000US-0214886.		
PR	30-JUN-2000; 2000US-0215135.		
PR	07-JUL-2000; 2000US-0216647.		
PR	07-JUL-2000; 2000US-0216880.		
PR	11-JUL-2000; 2000US-0217487.		
PR	11-JUL-2000; 2000US-0217496.		
PR	14-JUL-2000; 2000US-0218290.		
PR	26-JUL-2000; 2000US-0220963.		
PR	26-JUL-2000; 2000US-0220964.		
PR	14-AUG-2000; 2000US-0224518.		
PR	14-AUG-2000; 2000US-0224519.		
PR	14-AUG-2000; 2000US-0225213.		
PR	14-AUG-2000; 2000US-0225214.		
PR	14-AUG-2000; 2000US-0225266.		
PR	14-AUG-2000; 2000US-0225267.		
PR	14-AUG-2000; 2000US-0225268.		
PR	14-AUG-2000; 2000US-0225270.		
PR	14-AUG-2000; 2000US-0225447.		
PR	14-AUG-2000; 2000US-0225757.		
PR	14-AUG-2000; 2000US-0225758.		
PR	14-AUG-2000; 2000US-0225759.		
PR	18-AUG-2000; 2000US-0226279.		
PR	22-AUG-2000; 2000US-0226681.		
PR	22-AUG-2000; 2000US-0226688.		
PR	22-AUG-2000; 2000US-0227182.		
PR	23-AUG-2000; 2000US-0227009.		
PR			
PR	30-AUG-2000; 2000US-0228924.		
PR	01-SEP-2000; 2000US-0229287.		
PR	01-SEP-2000; 2000US-0229343.		
PR	01-SEP-2000; 2000US-0229344.		
PR	01-SEP-2000; 2000US-0229345.		
PR	05-SEP-2000; 2000US-0229509.		
PR	05-SEP-2000; 2000US-0229513.		
PR	06-SEP-2000; 2000US-0230437.		
PR	06-SEP-2000; 2000US-0230438.		
PR	08-SEP-2000; 2000US-0231242.		
PR	08-SEP-2000; 2000US-0231243.		
PR	08-SEP-2000; 2000US-0231244.		
PR	08-SEP-2000; 2000US-0231413.		
PR	08-SEP-2000; 2000US-0231414.		
PR	08-SEP-2000; 2000US-0232080.		
PR	08-SEP-2000; 2000US-0232081.		
PR	12-SEP-2000; 2000US-0231968.		
PR	14-SEP-2000; 2000US-0232397.		
PR	14-SEP-2000; 2000US-0232398.		
PR	14-SEP-2000; 2000US-0232399.		
PR	14-SEP-2000; 2000US-0232400.		
PR	14-SEP-2000; 2000US-0232401.		
PR	14-SEP-2000; 2000US-0233063.		
PR	14-SEP-2000; 2000US-0233064.		
PR	14-SEP-2000; 2000US-0233065.		
PR	21-SEP-2000; 2000US-0234223.		
PR	21-SEP-2000; 2000US-0234274.		
PR	25-SEP-2000; 2000US-0234997.		
PR	25-SEP-2000; 2000US-0234998.		
PR	26-SEP-2000; 2000US-0235484.		
PR	27-SEP-2000; 2000US-0235834.		
PR	27-SEP-2000; 2000US-0235836.		
PR	29-SEP-2000; 2000US-0236327.		
PR	29-SEP-2000; 2000US-0236367.		
PR	29-SEP-2000; 2000US-0236368.		
PR	29-SEP-2000; 2000US-0236369.		
PR	29-SEP-2000; 2000US-0236370.		
PR	02-OCT-2000; 2000US-0236802.		
PR	02-OCT-2000; 2000US-0237037.		
PR	02-OCT-2000; 2000US-0237038.		
PR	02-OCT-2000; 2000US-0237039.		
PR	02-OCT-2000; 2000US-0237040.		
PR	13-OCT-2000; 2000US-0239935.		
PR	13-OCT-2000; 2000US-0239937.		
PR	20-OCT-2000; 2000US-0240960.		
PR	20-OCT-2000; 2000US-0241221.		
PR	20-OCT-2000; 2000US-0241785.		
PR	20-OCT-2000; 2000US-0241786.		
PR	20-OCT-2000; 2000US-0241787.		
PR	20-OCT-2000; 2000US-0241808.		
PR	20-OCT-2000; 2000US-0241809.		
PR	20-OCT-2000; 2000US-0241826.		
PR	01-NOV-2000; 2000US-0244617.		
PR	08-NOV-2000; 2000US-0246474.		
PR	08-NOV-2000; 2000US-0246475.		
PR	08-NOV-2000; 2000US-0246476.		
PR	08-NOV-2000; 2000US-0246477.		
PR	08-NOV-2000; 2000US-0246478.		
PR	08-NOV-2000; 2000US-0246523.		
PR	08-NOV-2000; 2000US-0246524.		
PR	08-NOV-2000; 2000US-0246525.		
PR	08-NOV-2000; 2000US-0246526.		
PR	08-NOV-2000; 2000US-0246527.		
PR	08-NOV-2000; 2000US-0246528.		
PR	08-NOV-2000; 2000US-0246532.		
PR	08-NOV-2000; 2000US-0246609.		
PR	08-NOV-2000; 2000US-0246610.		
PR	08-NOV-2000; 2000US-0246611.		
PR	08-NOV-2000; 2000US-0246613.		
PR	17-NOV-2000; 2000US-0249207.		
PR	17-NOV-2000; 2000US-0249208.		
PR	17-NOV-2000; 2000US-0249209.		
PR	17-NOV-2000; 2000US-0249210.		

PR 17-NOV-2000: 2000US-0249211.
 PR 17-NOV-2000: 2000US-0249212.
 PR 17-NOV-2000: 2000US-0249213.
 PR 17-NOV-2000: 2000US-0249214.
 PR 17-NOV-2000: 2000US-0249215.
 PR 17-NOV-2000: 2000US-0249216.
 PR 17-NOV-2000: 2000US-0249217.
 PR 17-NOV-2000: 2000US-0249218.
 PR 17-NOV-2000: 2000US-0249244.
 PR 17-NOV-2000: 2000US-0249245.
 PR 17-NOV-2000: 2000US-0249264.
 PR 17-NOV-2000: 2000US-0249265.
 PR 17-NOV-2000: 2000US-0249297.
 PR 17-NOV-2000: 2000US-0249299.
 PR 17-NOV-2000: 2000US-0249300.
 PR 01-DEC-2000: 2000US-0250391.
 PR 05-DEC-2000: 2000US-0251030.
 PR 05-DEC-2000: 2000US-0251988.
 PR 05-DEC-2000: 2000US-0256719.
 PR 06-DEC-2000: 2000US-0251479.
 PR 08-DEC-2000: 2000US-0251856.
 PR 08-DEC-2000: 2000US-0251868.
 PR 08-DEC-2000: 2000US-0251869.
 PR 08-DEC-2000: 2000US-0251989.
 PR 11-DEC-2000: 2000US-0251990.
 PR 05-JAN-2001: 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.

XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI: 2001-465460/50.
 DR P-PSDB; AAU17541.
 XX

PT Novel polypeptides useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cancers, immune
 PT disorders and neuronal disorders -
 XX

PS Claim 1; SEQ ID No 493; 880pp; English.

XX The invention relates to novel isolated polypeptides (I), and
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
 CC diagnosing, preventing and treating diseases including immune system
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
 CC transplant rejections and graft versus host disease, infectious diseases
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative
 CC disorders, primary haematopoietic disorders, hyperproliferative
 CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
 CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
 CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
 CC Addison's disease), reproductive system disorders, gastrointestinal
 CC disorder (inflammatory disorders), liver disorders (cirrhosis),
 CC as stimulators of B-cell responsiveness to pathogens, activators of
 CC T-cells, to induce higher affinity antibodies, and as a means to induce
 CC tumour proliferation in pathologies e.g. acquired immune deficiency
 CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
 CC pathway protein coding sequences and PCR primers of the invention.
 XX

Alignment Scores:
 Pred. No.: 3,34e-126 Length: 1316
 Score: 1133.00 Matches: 220
 Percent Similarity: 99.10% Conservative: 0
 Best Local Similarity: 99.10% Mismatches: 2
 Query Match: 98.52% Indels: 0
 DB: 22 Gaps: 0

US-09-817-199A-2 (1-223) x AAS27458 (1-1316)
 QY 2 ThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerProPro 21
 |||||
 Db 1 ACGGGCAGCCAGGCGCGCTGGCCACCGGATGGCGAGGCCCGCGAGCGTCCCGCCC 60
 |||||
 QY 22 CysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyValGly 41
 |||||
 Db 61 TGCAGTCCGAGGTACGACCTCAGGGCAAGGTGATGCTTCTGGGAGACACAGCGCTCGGC 120
 |||||
 QY 42 LysThrCysPheLeuLeuGlnPheLysAspGlyValPheLeuSerGlyThrPheLeuAla 61
 |||||
 Db 121 AAACATGTTTCTGATCCATTCAAAGACGGGGCTTCTGTCGGAACCTTCATAGCC 180
 |||||
 QY 62 ThrValGlyLeuAspPheArgAsnLysValValThrValAspGlyValArgValLysLeu 81
 |||||
 Db 181 ACCGTCGGCATAGACTTCAGSACAAAGTGGTGACTGTGGATGGCGTGAGATGAAGCTG 240
 |||||
 QY 82 GlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArg 101
 |||||
 Db 241 CAGATCTGGGACACCGCTGGGAGGAACGTTCCGAAGCGTCAACCATGCTTATTACAGA 300
 |||||
 QY 102 AspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIle 121
 |||||
 Db 301 GATGCTCAGGCGCTTGCTTCTGCTGTATGACATCACCACAAATCTTCTTCGACAAATC 360
 |||||
 QY 122 ArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeuLeu 141
 |||||
 Db 361 AGGCGCTGGCTCACTCAGATTCTATGATATGCCAGAGGACGTGTGATCATGCTGCTA 420
 |||||
 QY 142 GlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeu 161
 |||||
 Db 421 GGCAACAAGCGGATATGAGCAGCAAGAGTATCGTTCGGAAGCGGAGAGACCTTG 480
 |||||
 QY 162 AlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGlu 181
 |||||
 Db 481 GCCAGGAGTACGGKGTCCCTTCTGAGACCGAGCGCAAGACTGGCATGAATGGAG 540
 |||||
 QY 182 LeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGlu 201
 |||||
 Db 541 TTAGCCTTCTGCCATCGCCAGGAACCTGAATACCGGGCGGCGATCAGCGGATGAG 600
 |||||
 QY 202 ProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysSer 221
 |||||
 Db 601 CCCAGCTTCAGATCCGAGACTATGTAKAGTCCAGAGAAGCGCTCCAGCTCTGCTCC 660
 |||||
 QY 222 PheMet 223
 |||||
 Db 661 TTCATG 666
 |||||
 RESULT 7
 AAX86720
 ID AAX86720 standard; DNA; 843 BP.
 AC AAX86720;
 XX
 DT 27-OCT-1999 . (first entry)
 XX
 DE DNA sequence encoding an exocytotic protein designated Exo2.
 XX
 KW Exocytotic protein; Exo2; exocytosis; Chediak-Higashi Syndrome;
 KW Inflammation; ss.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 124..699
 FT /tag= a
 FT /transl_except= (pos: 136..138, aa: Xaa)
 FT /transl_except= (pos: 202..204, aa: Xaa)
 FT /note= "Xaa is an unspecified amino acid"
 XX

PN W09942586-A2.
XX
PD 26-AUG-1999.
XX
PF 23-FEB-1999; 99WO-US03909.
XX
PR 26-MAY-1998; 98US-0086650.
PR 23-FEB-1998; 98US-0075534.
XX
XX (RIGE-) RIGEL PHARM INC.
XX
XX Anderson D, Fisher J, Huang B, Lorens J, Luo Y;
PI Shen M;
XX
XX WPI; 1999-518605/43.
DR P-PSDB; AAY30132.
XX
XX New exocytotic proteins useful for diagnosis and treatment of
PT exocytosis-mediated conditions and in drug screening
XX
XX Claim 6; Fig 6; 53pp; English.
XX
XX The present sequence encodes an exocytotic protein designated Exo2.
CC The protein has an inhibitory effect on exocytosis (i.e. the fusion of
CC secretory vesicles with the cellular plasma membrane) and can be
CC administered therapeutically to treat or prevent exocytosis-mediated
CC disorders e.g. Chediak-Higashi Syndrome (CHS) or conditions involving
CC undesirable release of compounds via exocytosis e.g. inflammation
CC mediated by the release of compounds such as histamine. Exo2 is useful
CC diagnostically and to produce antibodies useful to purify the proteins
CC and therapeutically to reduce or eliminate the biological activity of
CC the protein. Exo2 nucleic acids can be used therapeutically to increase
CC Exo2 activity in cells by known gene therapy techniques. They can also
CC be used to produce probes or primers to isolate Exo2 proteins from other
CC organisms, especially humans. The nucleic acids, host cells and proteins
CC are useful in screening assays to identify binding agents, especially
CC drug screening assays to identify agonists and antagonists useful
CC therapeutically to enhance or reduce Exo2 activity.
XX
XX Sequence 843 BP; 195 A; 218 C; 239 G; 175 T; 16 other;

Alignment Scores:
Pred. No.: 1.43e-114 Length: 843
Score: 1034.00 Matches: 205
Percent Similarity: 94.22% Conservative: 7
Best Local Similarity: 91.11% Mismatches: 11
Query Match: 89.91% Indels: 2
DB: 20 Gaps: 0

US-09-817-199A-2 (1-223) x AAX86720 (1-843)

Qy 1 MetThrGlyThrProGlyAla--ValAlaThrArgAspGlyClyAlaProGluArgSerP 20
Db 26 ATGAATGGCACACCGAGAGCTTGCCTGGGGATGGCGAGCGCCCTGAGCGCTCCC 85
Qy 20 roProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyV 40
Db 86 CGCCCTTCAGCCGCGAAGTACGATNTCCACGCGCAAGTGATGCTTCTGGAGANTCGGCG 145
Qy 40 aGlyLysThrCysPheLeuLeuGlnPheLysAspGlyAlaPheLeuSerGlyThrPheI 60
Db 146 TCGGCAAAACCTGTTTCTCGATCCAAATCAAGAGCGGGCTTCTGTCGGAACCTTNA 205
Qy 60 leAlaThrValGlyLeuPheArgAsnLysValValThrValAspGlyValArgValL 80
Db 206 TAGCCACCGTCGGCATACGATTCAGAAATAAGTGGTACAGTGGATGTCACGGGTGA 265
Qy 80 ysLeuGlnIleThrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrT 100
Db 266 AGCTTCAGATCTGGGACACTCGACGACAGGAGCGCTTCGCGAGTGTGACCCATGCTTAT 325
Qy 100 yrArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspA 120
XX

Db 326 ACCGAGATGCTCAGGCTTTGCTGCTGCTATGATGATCATCAACACGAGTCTCTTTTGACA 385
Qy 120 snlleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValIleMetL 140
XX
Db 386 ACATCAGGCGCTGGCTACAGAGATTCATGACTATGCCAGAGGACGCTGGTATTATGC 445
Qy 140 euLeuGlyAsnLysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluT 160
XX
Db 446 TTCTAGGCAACAAAGCCGATGTAGCAGCAAGAGGTGATCCGTTCTGAAGATGGAGAGA 505
Qy 160 hrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnV 180
XX
Db 506 CACTGGCCAGGGAATATGTTCTTTCATGGAGACCACTGCCAAGACTGGCATGAACG 565
Qy 180 alGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaA 200
XX
Db 566 TGGAGTTGGCTTCTGGCAATGGCAAGAACTGAATACCTGCAGGAGGACGCTG 625
Qy 200 spGluProSerPheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysC 220
XX
Db 626 ATGAGCCCGAGCTTCCAGATCCGAGACTATGTGGAGTCCCAAGAAGAGCGCTCCAGCTGCT 685
Qy 220 ysSerPheMet 223
Db 686 GCTCCTTTGTG 696
RESULT 8
ABA02774
ID ABA02774 standard; DNA; 576 BP.
XX
AC ABA02774;
XX
DT 07-FEB-2002 (first entry)
XX
DE Human degranulation regulator encoding DNA SEQ ID NO 3.
XX
KW Degranulation; mast cell; human; mouse; antiallergic; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..576
FT /*tag= a /product= "degranulation regulator"
XX
PN WO200179478-A1.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-JP03268.
XX
PR 19-APR-2000; 2000JP-0118408.
XX
PA (DAIN) DAINIPPON PHARM CO LTD.
XX
PI Yamada T, Ido M;
XX
XX WPI; 2002-041335/05.
DR P-PSDB; AAM52190.
XX
PT Mast cell degranulation controller for treatment of allergies -
XX
PS Claim 11; Page 64-65; 85pp; Japanese.
XX
CC The invention relates to a protein for regulating degranulation of mast
CC cells (degranulation regulators) and the encoding polynucleotides, with
CC antiallergic activity, used in the treatment of allergies associated
CC with degranulation of mast cells.
XX
SQ Sequence 576 BP; 138 A; 147 C; 168 G; 123 T; 0 other;
Alignment Scores:
Pred. No.: 5.49e-107 Length: 576

Score: 969.00 Matches: 189
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 98.95% Mismatches: 0
Query Match: 84.26% Indels: 0
DB: 24 Gaps: 0

US-09-817-199A-2 (1-223) x ABA02774 (1-576)

```
QY 33 MetLeuLeuGlyAspThrGlyValGlyLysThrCysPheLeuLeuGlnPheLysAspGly 52
DB 1 ATGCTTCTTGGAGACTCGGCGCTCGCAAAACATGTTCTCTGATCCAAATCAAGACGGG 60
QY 53 AlaPheLeuSerGlyThrPheIleAlaThrValGlyIleAspPheArgAnLysValVal 72
DB 61 GCCTTCTCTCGGAACCTTATAGCACCGTCGGCATAGACTTCAGGAACAAGGTGGTG 120
QY 73 ThrValAspGlyValArgValLysLeuGlnIleTyrAspThrAlaGlyGlnGluArgPhe 92
DB 121 ACTGTGGATGGGTGAGATGAGCTGACAGATCTGGGACCCCTGGCGAAGACGGTTC 180
QY 93 ArgSerValThrHisAlaTyrTyrArgAspAlaGlnAlaLeuLeuLeuTyrAspIle 112
DB 181 CGAAGCGTCACCATCTTATTACAGAGATGCTCAGGCGCTTGCTTCTGTGTATGACATC 240
QY 113 ThrAsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAla 132
DB 241 ACCAACAAATCTCTTTCACACACATCAGGGCGCTGGCTCAGTATTCATGATGATGCC 300
QY 133 GlnArgAspValIleMetLeuLeuGlyAsnLysAlaAspMetSerSerGluArgVal 152
DB 301 CAGAGGAGCGTGGTGATCATGCTGTAGGCACACAGCGCGATATGAGCAGCAAGAGTG 360
QY 153 IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrGlyValProPheLeuGluThr 172
DB 361 ATCCGTTCCGAAGACGAGACCTTGGCCAGGAGTACGGTGTCTCTCTGGAGACC 420
QY 173 SerAlaLysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys 192
DB 421 AGCGCCAAAGACTGGCATGAATGTGGAGTTAGGCTTCTTGCCCATCCCAAGAACTGAAA 480
QY 193 TyrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer 212
DB 481 TACCGGCGCGGCATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 540
```

RESULT 9

ABA02773

ID ABA02773 standard; DNA; 576 BP.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX

PR

PA

XX

XX

PI

XX

DR

DR

XX

XX

PT

PS

CC

CC

CC

CC

XX

SQ

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-817-199A-2 (1-223) x ABA02773 (1-576)

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

19-APR-2000; 2000JP-0118408.

(DAIN) DAINIPPON PHARM CO LTD.

Yamada T, Ido M;

WPI; 2002-041335/05.

P-PSDB; AAM52189.

Mast cell degranulation controller for treatment of allergies

Claim 10; Page 63-64; 85pp; Japanese.

The invention relates to a protein for regulating degranulation of mast

cells (degranulation regulators) and the encoding polynucleotides, with

antiallergic activity, used in the treatment of allergies associated

with degranulation of mast cells.

Sequence 576 BP; 139 A; 139 C; 165 G; 133 T; 0 other;

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-817-199A-2 (1-223) x ABA02773 (1-576)

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

19-APR-2000; 2000JP-0118408.

(DAIN) DAINIPPON PHARM CO LTD.

Yamada T, Ido M;

WPI; 2002-041335/05.

P-PSDB; AAM52189.

Mast cell degranulation controller for treatment of allergies

Claim 10; Page 63-64; 85pp; Japanese.

The invention relates to a protein for regulating degranulation of mast

cells (degranulation regulators) and the encoding polynucleotides, with

antiallergic activity, used in the treatment of allergies associated

with degranulation of mast cells.

Sequence 576 BP; 139 A; 139 C; 165 G; 133 T; 0 other;

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-817-199A-2 (1-223) x ABA02773 (1-576)

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

19-APR-2000; 2000JP-0118408.

(DAIN) DAINIPPON PHARM CO LTD.

Yamada T, Ido M;

WPI; 2002-041335/05.

P-PSDB; AAM52189.

Mast cell degranulation controller for treatment of allergies

Claim 10; Page 63-64; 85pp; Japanese.

The invention relates to a protein for regulating degranulation of mast

cells (degranulation regulators) and the encoding polynucleotides, with

antiallergic activity, used in the treatment of allergies associated

with degranulation of mast cells.

Sequence 576 BP; 139 A; 139 C; 165 G; 133 T; 0 other;

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-817-199A-2 (1-223) x ABA02773 (1-576)

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

AA881431 standard; cDNA; 475 BP.
AA881431;
13-FEB-2002 (first entry)
DNA encoding novel human diagnostic protein #17235.
Human; chromosome mapping; gene mapping; gene therapy; forensic;
food supplement; medical imaging; diagnostic; genetic disorder; ss.
Homo sapiens.
WO200175067-A2.
11-OCT-2001.
30-MAR-2001; 2001WO-US08631.
31-MAR-2000; 2000US-0540217.
23-AUG-2000; 2000US-0649167.
(HYSE-) HYSEQ INC.
Drmanac RT, Liu C, Tang YT;
WPI; 2001-639362/73.
P-PSDB; ABG17244.
New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity -
Claim 1; SEQ ID No 17235; 103pp; English.
The invention relates to isolated polynucleotide (I) and
polypeptide (II) sequences. (I) is useful as hybridisation probes,
polymerase chain reaction (PCR) primers, oligomers, and for chromosome
and gene mapping, and in recombinant production of (II). The
polynucleotides are also used in diagnostics as expressed sequence tags
for identifying expressed genes. (I) is useful in gene therapy techniques
to restore normal activity of (II) or to treat disease states involving
(II). (II) is useful for generating antibodies against it, detecting or
quantitating a polypeptide in tissue, as molecular weight markers and as
a food supplement. (II) and its binding partners are useful in medical
imaging of sites expressing (II). (I) and (II) are useful for treating
disorders involving aberrant protein expression or biological activity.
The polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits to assess biodiversity
and to produce other types of data and products dependent on DNA and
amino acid sequences. AA881431-17235 represent novel human
diagnostic coding sequences of the invention.
Note: The sequence data for this patent did not appear in the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.
Sequence 475 BP; 108 A; 136 C; 138 G; 93 T; 0 other;
Alignment Scores:
Pred. No.: 3.58e-83 Length: 475
Score: 770.00 Matches: 148
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.33% Mismatches: 0
Query Match: 66.96% Indels: 0
DB: 23 Gaps: 0
US-09-817-199A-2 (1-223) x AA881431 (1-475)
Qy 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
Db 28 ATGACGGGACGCCAGGCGCGTTCACCGCGGATGCGAGGCCCGCGCGCTCCCG 87

Qy 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
AC |||||||
Db 88 CCCTGCAGTCCGAGCTACGACCTCAGCGCAAGGTGATGCTTCTGGGACACACGGCGTC 147
Qy 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
KW |||||||
Db 148 GGCACAAACATGTTCTCTGATCCCAATTCAAGACGGGGCTTCTCTGCCGAACCTTCATA 207
Qy 61 AlaThrValGlyIleAspPheArgAsnLysValThrValAspGlyValArgValLys 80
KW |||||||
Db 208 GCCACCGCGGCATAGACTTCAGGAACAAGTGGTGACTGTGGATGGCGTGAAGTGAAG 267
Qy 81 LeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
KW |||||||
Db 268 CTGCAGATCTGGGACACCGCTGGCGCAGGAACGGTCCGAAGCGTCACCCATCTATTAC 327
Qy 101 ArgAspAlaGlnAlaLeuLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsn 120
KW |||||||
Db 328 AGAGATGCTCAGGCGCTTCTGCTGTATGATCATCACCACAAATCTTCTTTTCGACAC 387
Qy 121 IleArgAlaTrpLeuThrClnIleHisGluTyrAlaGlnArgAspValIleMetLeu 140
KW |||||||
Db 388 ATCAGGGCGTGGCTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 447
Qy 141 LeuGlyAsnLysAlaAspMetSerSer 149
KW |||||||
Db 448 CTAGGCAACAGCGCGGATATGAGCAGC 474
RESULT 11
AAV65197
ID AAV65197 standard; DNA; 1340 BP.
XX AAV65197;
XX AC AAV65197;
XX DT 17-DEC-1998 (first entry)
XX DE Human RAB protein, SRAB, coding sequence.
XX KW Human; RAB protein; SRAB; vesicle trafficking disorder; epilepsy; cancer;
KW Huntington's disease; Parkinson's disease; schizophrenia; choroideraemia;
KW viral infection; therapy; autoimmune sialosis; cystic fibrosis;
KW diabetes mellitus; hyperglycaemia; tubulointerstitial nephritis;
KW hypoglycaemia; pancreatic enzyme deficiency; pancreatitis;
KW polycystic renal disease; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
CDS 43..615
FT /*tag= a
FT /product= "SRAB"
XX FT
XX PN WO9842839-A1.
XX PD 01-OCT-1998.
XX PF 25-MAR-1998; 98WO-US05871.
XX PR 26-MAR-1997; 97US-0824873.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Guegler KJ, Hillman JL;
XX DR WPI; 1998-54282/46.
XX DR P-PSDB; AA880747.
XX PT New human Rab protein, SRAB - useful for treating disorders
PT associated with SRAB expression, including vesicle trafficking,
XX viral infection, and cancer
PS Claim 5; Fig 1; 31pp; English.

D	b	1477	TATGTGGCGGAGCTATTTCCTCAGCACACGTTGGCAATGTATTTTAGGAACAATAAGTGGTGTC	1536
Q	y	74	ValAspGlyValArgValLysLeuGlnIleTrpAspThrAlaGlyGlnGluArgPheArg	93
D	b	1537	GTCGATGAACGC CGCTCAACGCTGCAAAATCTGGGACACAGCTGGTCAGGAGCATCCGG	1596
Q	y	94	SerValThrHisAlaTyrrTyrrArgaspAlaGlnAlaLeuLeuLeuLeuTyrrAspIleThr	113
D	b	1597	AGCGTTACCACGCGCTATTATTCGGACGCGCACGCTACTGCTGCTGTACGACGTGACC	1656
Q	y	114	AsnLysSerSerPheAspAsnIleArgAlaTrpLeuThrGluIleHisGluTyrAlaGln	133
D	b	1657	AACAAGACCACCTATGACAACATTCGCCGCTGGCTGGGCGAGATCCGGAGTACGCGCAG	1716
Q	y	134	ArgAspValValIleMetLeuLeuGlyAsnLysAlaAspMetSer---SerGluArgVal	152
D	b	1717	GAGCAGCTGGTCATCGTTTTTAATAGGCAACAAGCGCGACTGCAGCGGCAGCGCGCAG	1776
Q	y	153	IleArgSerGluAspGlyGluThrLeuAlaArgGluTyrrGlyValProPheLeuGluThr	172
D	b	1777	GTGAAGCGGAGGATGGGAGCGTTTGGGCGGGAGCACAACGTCGCCCTTCATGGAGACC	1836
Q	y	173	SerAlalysThrGlyMetAsnValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLys	192
D	b	1837	TCGCCCAAGACGGGACTCAATGTGGAGTGTCTTCACAGCGGTGGCCAGGCAACTAAAG	1896
Q	y	193	TyrrArgAlaGlyHisGlnAlaAspGluProSerPheGlnIleArgAspTyrValGluSer	212
D	b	1897	AGTCGGCGCTACGAGCAGCGCGATGATGGAAAGTTCATGTGTCATGATTTTGTGCGTGAC	1956
Q	y	213	GlnLysLysArgSerSerCysCysSer	221
D	b	1957	AATCAAAGCGCGCTCGGTTTTCGCC	1983
RESULT 13				
ID	AAS27040 standard; cDNA; 964 BP.			
XX	AAS27040;			
XX				
XX	(first entry)			
DE	07-NOV-2001			
XX	cDNA encoding novel signal transduction pathway protein, Seq ID 75.			
KW	Neuroprotective; cytostatic; dermatologic; immunosuppressive; cancer;			
KW	antiinflammatory; anti-HIV; antibacterial; antineoplastic; tumor;			
KW	immune system disorder; rheumatoid arthritis; inflammatory condition;			
KW	organ transplant rejection; infection; hepatitis C; blood disorder;			
KW	sickle cell anemia; hyperproliferative disorder; Gaucher's disease;			
KW	chromosomal abnormality; Down syndrome; ischaemia; renal disorder;			
KW	cardiovascular; respiratory; wound healing; endocrine; Addison's disease;			
KW	reproductive system; gastrointestinal; liver disorder; AIDS; ss;			
XX	acquired immune deficiency syndrome.			
XX	Homo sapiens.			
XX	WO200154733-A1.			
P	02-AUG-2001.			
XX	17-JAN-2001; 2001WO-US01312.			
XX	31-JAN-2000; 2000US-0179065.			
PR	04-FEB-2000; 2000US-0180628.			
PR	24-FEB-2000; 2000US-0184664.			
PR	16-MAR-2000; 2000US-0186350.			
PR	02-MAR-2000; 2000US-0189874.			
PR	17-MAR-2000; 2000US-0190076.			
PR	18-APR-2000; 2000US-0198123.			
PR	19-MAY-2000; 2000US-0205515.			
PR	27-JUN-2000; 2000US-0209467.			
PR	28-JUN-2000; 2000US-0214886.			

```
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0248475.
PR 08-NOV-2000; 2000US-0248476.
PR 08-NOV-2000; 2000US-0248477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0248525.
PR 08-NOV-2000; 2000US-0248526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0248610.
PR 08-NOV-2000; 2000US-0248611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0246617.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI: 2001-465460/50.
DR P-PSDB; AAU17123.
XX
PT Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders -
XX
XX Claim 1; SEQ ID No 75; 880pp; English.
PS
XX The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC
```

(e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and other blood-related disorders (sickle cell anaemia), myeloproliferative disorders, primary haematopoietic disorders, hyperproliferative disorders (e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g. glomerulonephritis), cardiovascular disorders (e.g. arrhythmia), respiratory disorders, dermatological disorders, in wound healing, epithelial cell proliferation, endocrine disorders (e.g. Addison's disease), reproductive system disorders, gastrointestinal disorder (inflammatory disorders), liver disorders (cirrhosis), as stimulators of B-cell responsiveness to pathogens, activators of T-cells, to induce higher affinity antibodies, and as a means to induce tumour proliferation in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction CC pathway protein coding sequences and PCR primers of the invention. XX

Alignment Scores:
Pred. No.: 1.79e-54 Length: 964
Score: 535.00 Matches: 98
Percent Similarity: 85.40% Conservative: 19
Best Local Similarity: 71.53% Mismatches: 20
Query Match: 46.52% Indels: 0
DB: 22 Gaps: 0

US-09-817-199A-2 (1-223) x AAS27040 (1-964)

QY 84 TrpAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyrArgAspAla 103
|||||
DB 3 TGGGACACAGCTGGTCAGAGCGGTTCGGCAGTGTTACCATGCTACTACCGGGATGCT 62
|||||
QY 104 GlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerSerPheAspAsnIleArgAla 123
|||||
DB 63 CATGCTCTGCTGCTCTACGATGTCACCAACAGGCTCCTTTGACAAATCCAGGCC 122
|||||
QY 124 TrpLeuThrGluIleHisGluTyrAlaGlnArgAspValIleMetLeuLeuGlyAsn 143
|||||
DB 123 TGGCTGACCGAGATCCACGAGTACGCCACGACGACGCTGCGCTCATGCTGTGGGAAC 182
|||||
QY 144 LysAlaAspMetSerSerGluArgValIleArgSerGluAspGlyGluThrLeuAlaArg 163
|||||
DB 183 AAGGTGGACTCTGCCATGAGCGTGTGGTAAGAGGAGGACGGGAGAGCTGGGCCAAG 242
|||||
QY 164 GluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsnValGluLeuAla 183
|||||
DB 243 GAGTATGGACTGCCCTTCATGGAGACCGCCCAACAGCGGCTCAACGTGGACTGGCC 302
|||||
QY 184 PheLeuAlaIleAlaLysGluLeuLysTyrArgAlaGlyHisGlnAlaAspGluProSer 203
|||||
DB 303 TTCACAGCCATAGCAAAAGAGTTGAAGCAGCGCTCCATGAAGGCTCCCGAGCGCGCGC 362
|||||
QY 204 PheGlnIleArgAspTyrValGluSerGlnLysLysArgSerSerCysCys 220
|||||
DB 363 TTCCGGCTGCATGATTACGTTAAGAGGAGGAGGTCTAGAGGGGCTCTCTGCTGC 413
|||||

RESULT 14
AAC38429
ID AAC38429 standard; DNA; 911 BP.
XX
AC AAC38429;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 20938.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.

XX	06-SEP-2000.	PR	02-JUL-1999;	99US-0142055.
PD		PR	06-JUL-1999;	99US-0142390.
XX		PR	08-JUL-1999;	99US-0142803.
PF		PR	09-JUL-1999;	99US-0142920.
XX	25-FEB-2000; 2000EP-0301439.	PR	12-JUL-1999;	99US-0142977.
PR	25-FEB-1999;	PR	13-JUL-1999;	99US-0143542.
PR	05-MAR-1999;	PR	14-JUL-1999;	99US-0143624.
PR	09-MAR-1999;	PR	15-JUL-1999;	99US-0144005.
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144085.
PR	25-MAR-1999;	PR	16-JUL-1999;	99US-0144086.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144325.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144331.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144332.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144333.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144334.
PR	19-APR-1999;	PR	19-JUL-1999;	99US-0144335.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144335.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144352.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	28-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0144884.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145086.
PR	04-MAY-1999;	PR	21-JUL-1999;	99US-0145088.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145085.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145089.
PR	07-MAY-1999;	PR	22-JUL-1999;	99US-0145192.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145145.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913.
PR	18-MAY-1999;	PR	27-JUL-1999;	99US-0145918.
PR	19-MAY-1999;	PR	27-JUL-1999;	99US-0145919.
PR	20-MAY-1999;	PR	28-JUL-1999;	99US-0145951.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146386.
PR	24-MAY-1999;	PR	02-AUG-1999;	99US-0146388.
PR	25-MAY-1999;	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	PR	03-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171.
PR	10-JUN-1999;	PR	11-AUG-1999;	99US-0148319.
PR	14-JUN-1999;	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065.
PR	21-JUN-1999;	PR	27-AUG-1999;	99US-0151066.
PR	22-JUN-1999;	PR	30-AUG-1999;	99US-0151080.
PR	23-JUN-1999;	PR	31-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	PR	01-SEP-1999;	99US-0151338.
PR	24-JUN-1999;	PR	07-SEP-1999;	99US-0151930.
PR	28-JUN-1999;	PR	10-SEP-1999;	99US-0152363.
PR	29-JUN-1999;	PR	13-SEP-1999;	99US-0153070.
PR	30-JUN-1999;	PR	13-SEP-1999;	99US-0153758.
PR	01-JUL-1999;	PR	16-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	PR	20-SEP-1999;	99US-0154039.
PR		PR		99US-0154779.

PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.: 3,18e-52 Length: 911
Score: 516.00 Matches: 107
Percent Similarity: 66.52% Conservative: 48
Best Local Similarity: 45.92% Mismatches: 58
Query Match: 44.87% Indels: 20
DB: 21 Gaps: 6

US-09-817-199a-2 (1-223) x AAC38429 (1-911)

Qy 1 MetThrGlyThrProGlyAlaValAlaThrArgAspGlyGluAlaProGluArgSerPro 20
Db 28 CTAACATCTTCTCGATGCGGTT-----GCGCGCGCAAGAGCT--- 66
Qy 21 ProCysSerProSerTyrAspLeuThrGlyLysValMetLeuLeuGlyAspThrGlyVal 40
Db 67 -----CGTTACAGACTATGATTATCTCATCAAGCTTCTCTCATCGGTATAGCGGTG 120
Qy 41 GlyLysThrCysPheLeuIleGlnPheLysAspGlyAlaPheLeuSerGlyThrPheIle 60
Db 121 GAAAAAGTCTCTGTTACTCGGTTTTTCGATGATACTTTCACCAGC---AGTTTCATT 177
Qy 61 AlaThrValGlyIleAspPheArgAsnLysValValThrValAspGlyValArgValLys 80
Db 178 ACTACCATGGAATCGACTTCAAGATAAGAACAGCTCGAAGCTTGATGGAAGCGTATCAAA 237
Qy 81 LeuGlnIleThrAspThrAlaGlyGlnGluArgPheArgSerValThrHisAlaTyrTyr 100
Db 238 TTGCAGATATGGGACACTGCTGGSACAAGCGTTTATAGACTATAACACAGCATATTAC 297
Qy 101 ArgAspAlaGlnAlaLeuLeuLeuTyrAspIleThrAsnLysSerPheAspAsn 120
Db 298 AGAGGAGCGATGGGTATATTACTCGTCTATGATATACAGATGAGTCATCATCAACAAC 357

Qy 121 IleArgAlaTrpLeuThrGluIleHisGluTyrAlaGlnArgAspValValIleMetLeu 140
Db 358 ATTAGGAAGTGGATGAAAAACATTGAGCAGCATGCGTCGATAGTGTCAACAAGATATTG 417
Qy 141 LeuGlyAsnLysAlaAspMet---SerSerGluArgValIleArgSerGluAspGlyGlu 159
Db 418 GTCGGTAACAAGCGCAGCATGAGCAAGAAAGGGGTGTCACCAACATCAAAAGGACAA 477
Qy 160 ThrLeuAlaArgGluTyrGlyValProPheLeuGluThrSerAlaLysThrGlyMetAsn 179
Db 478 GCTCTTGCTGATGATGATGGAATCAATCTTTGAGAGCAGTGCACAAACACCAAGAA 537
Qy 180 ValGluLeuAlaPheLeuAlaIleAlaLysGluLeuLysTyrArg-----AlaGly 196
Db 538 GTCGAGCAGGTTTCTCTATCGCTAAAGACATAAAACAAAGACTCACAGAAAGCGAT 597
Qy 197 HisGlnAlaAspGluProSerPheGlnIle-----ArgAspTyrVal 210
Db 598 ACAAAAGCGGAGCCCAAGGGATCAAGATCACTAAACAAGATGCTAACAAAGCCTCATCG 657
Qy 211 GluSerGlnLysLysArgSerSerCysCysSerPheMet 223
Db 658 TCCTTCTACAAATGAGAAATCAGCTTGTCTGCTGCTTATGTT 696
RESULT 15
AAC44482
ID AAC44482 standard; DNA; 1129 BP.
AC AAC44482;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays DNA fragment SEQ ID NO: 42986.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.
OS Zea mays subsp. mays.
PN EP1033405-A2.
PD 06-SEP-2000.
PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2002, 16:39:28 ; Search time 20 seconds
(without alignments)
167.929 Million cell updates/sec

Title: US-09-817-199A-2

Perfect score: 1150
Sequence: 1 MTGTPGAVATRDGEAPERSP.....FOIRDYVESQKRSSCCSFM 223

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 97044 seqs, 15060890 residues

Total number of hits satisfying chosen parameters: 97044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	1150	100.0	223	10	US-09-817-199A-2
2	1081	94.0	223	10	US-09-817-199A-4
3	977	85.0	191	10	US-09-794-257-14
4	977	85.0	191	12	US-10-051-986-3
5	547	47.6	106	10	US-09-867-550-1812
6	501	43.6	207	10	US-09-794-257-8
7	500	43.5	190	10	US-09-822-860-5
8	498	43.3	218	10	US-09-925-300-1571
9	497.5	43.3	162	10	US-09-834-765-766
10	483	42.0	245	10	US-09-925-302-534
11	472	41.0	201	10	US-09-967-736-8
12	468	40.7	201	10	US-09-967-736-3
13	466	40.5	198	10	US-09-794-257-16
14	466	40.5	198	10	US-09-945-173-5
15	466	40.5	198	10	US-09-972-529-4
16	458.5	39.9	222	10	US-09-820-003A-4
17	448.5	39.0	212	10	US-09-350-874-67
18	429	37.3	212	10	US-09-817-198A-2
19	427.5	37.2	212	10	US-09-817-198A-4

20	416	36.2	218	10	US-09-817-198A-5
21	413.5	36.0	312	10	US-09-925-302-783
22	386	33.6	832	10	US-09-834-765-2
23	373.5	32.5	213	10	US-09-794-257-5
24	365	31.7	208	9	US-10-108-605-45
25	364.5	31.7	624	10	US-09-834-765-5
26	364.5	31.7	625	10	US-09-834-765-762
27	360.5	31.3	168	10	US-09-834-765-765
28	360	31.3	239	10	US-09-925-301-1077
29	359	31.2	217	10	US-09-988-974-3
30	358.5	31.2	213	10	US-09-988-974-8
31	352.5	30.7	217	10	US-09-925-300-1364
32	350	30.4	201	10	US-09-822-860-2
33	349.5	30.4	161	10	US-09-834-765-763
34	344.5	30.0	211	12	US-10-051-986-6
35	339	29.5	216	10	US-09-945-173-10
36	319.5	27.8	157	10	US-09-834-765-764
37	314.5	27.3	173	10	US-09-820-003A-2
38	313.5	27.3	206	10	US-09-828-310-13
39	305.5	26.6	201	10	US-09-988-974-9
40	302.5	26.3	201	10	US-09-988-974-5
41	302.5	26.3	209	10	US-09-864-761-42996
42	299.5	26.0	208	10	US-09-925-302-629
43	294.5	25.6	137	10	US-09-817-182-6
44	284.5	24.7	259	12	US-10-051-986-1
45	284.5	24.7	260	12	US-10-051-986-4

ALIGNMENTS

RESULT 1

US-09-817-199A-2

; Sequence 2, Application US/09817199A

; Patent No. US20(2014)2380A1

; GENERAL INFORMATION:

; APPLICANT: SHAO, Wei et al.

; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE

; TITLE OF INVENTION: PROTEINS, AND USES THEREOF

; FILE REFERENCE: CL001187

; CURRENT APPLICATION NUMBER: US/09/817,199A

; CURRENT FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 223

; TYPE: PRT

; ORGANISM: Human

US-09-817-199A-2

Query Match 100.0%; Score 1150; DB 10; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.5e-116;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI	60
DB	1	MTGTPGAVATRDGEAPERSPPCSPSYDLTGKVMLLGDTGVGKTCFLIQFKDGAFLSGTFI	60
QY	61	ATVGDIDFNKVVTVDGVVRVKLQIWDTAGQERFSVTHAYYRDAQALLLYDITNKSSFDN	120
DB	61	ATVGDIDFNKVVTVDGVVRVKLQIWDTAGQERFSVTHAYYRDAQALLLYDITNKSSFDN	120
QY	121	IRAWLTIHEYAQRDVIMLLGNKADMSSEVIRSEDGETLAREYGVPLETSAKTGMNV	180
DB	121	IRAWLTIHEYAQRDVIMLLGNKADMSSEVIRSEDGETLAREYGVPLETSAKTGMNV	180
QY	181	ELAFIAIAKELKYRAGHQADEPSFOIRDYVESQKRSSCCSFM	223
DB	181	ELAFIAIAKELKYRAGHQADEPSFOIRDYVESQKRSSCCSFM	223

RESULT 2

US-09-817-199a-4
; Sequence 4, Application US/09817199a
; Patent No. US20020142380A1
; GENERAL INFORMATION:
; APPLICANT: SHAO, Wei et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001187
; CURRENT APPLICATION NUMBER: US/09/817,199a
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-817-199a-4

Query Match 94.0%; Score 1081; DB 10; Length 223;
Best Local Similarity 93.7%; Pred. No. 4.2e-109;
Matches 209; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
QY 1 MCTPGCAVTRDGEAPERSPPSPDLTGKVMLLGDTGVGKTCFLIQKDGAFLSGTFI 60
Db 1 MCTPGCAATAGDGEAPERSPPSPDLTGKVMLLGDSGVGKTCFLIQKDGAFLSGTFI 60
QY 61 ATVGIDFRNKVTVDGVVRVKLQIWDTAGOERFSVTHAYYRDAQAALLLYDITNKSSFDN 120
Db 61 ATVGIDFRNKVTVDGVARVKLQIWDTAGOERFSVTHAYYRDAQAALLLYDITNKSSFDN 120
QY 121 IRALWTEIHEYAQRDVVIMLLGNKADMSSEVIRSEDGETLAREYGVFPFLETSKATGMNV 180
Db 121 IRALWTEIHEYAQRDVVIMLLGNKADVSSERVIRSEDGETLAREYGVFPFLETSKATGMNV 180
QY 181 ELAFLAIKELKYRAGHQADEFSQTRDYVESQKRRSSCCSFM 223
Db 181 ELAFLAIKELKYRAGROPDEFSQIRDYVESQKRRSSCCSFEV 223

RESULT 3
US-09-794-257-14
; Sequence 14, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1e1
; FILE REFERENCE: 35800/209285
; CURRENT APPLICATION NUMBER: US/09/794,257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 191
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-794-257-14

Query Match 85.0%; Score 977; DB 10; Length 191;
Best Local Similarity 100.0%; Pred. No. 5.5e-98;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 33 MLIGDGTGVGKTCFLIQKDGAFLSGTFTIATVGIDFRNKVTVDGVVRVKLQIWDTAGOERF 92
Db 1 MLIGDGTGVGKTCFLIQKDGAFLSGTFTIATVGIDFRNKVTVDGVVRVKLQIWDTAGOERF 60
QY 93 RSVTHAYYRDAQAALLLYDITNKSSFDNIRALWTEIHEYAQRDVVIMLLGNKADMSSEV 152
Db 61 RSVTHAYYRDAQAALLLYDITNKSSFDNIRALWTEIHEYAQRDVVIMLLGNKADMSSEV 120

QY 153 IRSEDGETLAREYGVFPFLETSKATGMNVFLAFLAIKELKYRAGHQADEFSQIRDYVES 212
Db 121 IRSEDGETLAREYGVFPFLETSKATGMNVFLAFLAIKELKYRAGHQADEFSQIRDYVES 180
QY 213 QKRRSSCCSFM 223
Db 181 QKRRSSCCSFM 191
RESULT 4
US-10-051-986-3
; Sequence 3, Application US/10051986
; Patent No. US20020146770A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; Tang, Y. Tom
; Lal, Preeti
; Guegler, Karl J.
; Corley, Neil C.
; Patterson, Chandra
; Batra, Sajeev
; Baughn, Mariah R.
; TITLE OF INVENTION: RAS PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/051,986
; FILING DATE: 15-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/766,551
; FILING DATE: DECEMBER 12, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0168-1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: UCMCL5T01
; CLONE: 1528559
; SEQUENCE DESCRIPTION: SEQ ID NO: 3 :
US-10-051-986-3

Query Match 85.0%; Score 977; DB 12; Length 191;
Best Local Similarity 100.0%; Pred. No. 5.5e-98;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 33 MLIGDGTGVGKTCFLIQKDGAFLSGTFTIATVGIDFRNKVTVDGVVRVKLQIWDTAGOERF 92
Db 1 MLIGDGTGVGKTCFLIQKDGAFLSGTFTIATVGIDFRNKVTVDGVVRVKLQIWDTAGOERF 60
QY 93 RSVTHAYYRDAQAALLLYDITNKSSFDNIRALWTEIHEYAQRDVVIMLLGNKADMSSEV 152

Db 61 RSVTHAYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLGNKADSSERV 120

Qy 153 IRSEDETLAREYGVPFLETSAKTGMNVELAFALAKELKYRAGHQADPSFOIRDYVES 212

Db 121 IRSEDETLAREYGVPFLETSAKTGMNVELAFALAKELKYRAGHQADPSFOIRDYVES 180

Qy 213 QKRRSSCCSFM 223

Db 181 QKRRSSCCSFM 191

RESULT 5

US-09-867-550-1812

; Sequence 1812, Application US/09867550

; Patent No. US20020082206A1

; GENERAL INFORMATION:

; APPLICANT: Leach, Martin D.

; APPLICANT: Mehraban, Fuad,

; APPLICANT: Conley, Pamela

; APPLICANT: Law, Debbie

; APPLICANT: Topper, James

; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and

; TITLE OF INVENTION: Thereby

; FILE REFERENCE: 21402-013 (Cura-313)

; CURRENT APPLICATION NUMBER: US/09/867,550

; CURRENT FILING DATE: 2001-09-20

; PRIOR APPLICATION NUMBER: USSN 60/208,427

; PRIOR FILING DATE: 2000-05-30

; NUMBER OF SEQ ID NOS: 2125

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1812

; LENGTH: 106

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-867-550-1812

Query Match 47.6%; Score 547; DB 10; Length 106;

Best Local Similarity 100.0%; Pred. No. 5.5e-52;

Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 MLIGDTGKTCFLIOFKDGAFLSGTFTATVGIDFRNKVTVVDGVRVKLIQIWDTAGQERF 92

Db 1 MLIGDTGKTCFLIOFKDGAFLSGTFTATVGIDFRNKVTVVDGVRVKLIQIWDTAGQERF 60

Qy 93 RSVTHAYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVI 138

Db 61 RSVTHAYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVI 106

RESULT 6

US-09-794-257-8

; Sequence 8, Application US/09794257

; Patent No. US2002009804A1

; GENERAL INFORMATION:

; APPLICANT: Meyers, Rachel

; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US2002009804A1el

; TITLE OF INVENTION: Human G-Proteins

; FILE REFERENCE: 35800/209285

; CURRENT APPLICATION NUMBER: US/09/794,257

; CURRENT FILING DATE: 2001-02-27

; PRIOR APPLICATION NUMBER: 60/185,606

; PRIOR FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 207

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-794-257-8

Query Match 43.6%; Score 501; DB 10; Length 207;

Best Local Similarity 49.2%; Pred. No. 1.3e-46;

Matches 98; Conservative 40; Mismatches 59; Indels 2; Gaps 2;

Qy 25 SYDLGCKVMLLGDGTGKTCFLIOFKDGAFLSGTFTATVGIDFRNKVTVVDGVRVKLIQIW 84

Db 4 TYDLFKLLIGDSGVKTCCLFRSEDAF-NTTISTIGIDFKIRTIELDGKKIKLQIW 62

Qy 85 DTAGQEFERSVTHAYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLGNK 144

Db 63 DTAGQEFERTITAYRGAMGIMLYDITNEKSFNKNWIRNIEEHASDVVERMILGNK 122

Qy 145 ADMSEFVIRSEDETLAREYGVPFLETSAKTGMNVELAFALAKELKYRAGHQADPSFOIR 203

Db 123 CDMDKRVQSKERGEKLAIDYIGIKFLETSAKSSANVEEAFFTLARDIMTKLNKKNDSNS 182

Qy 204 FOIRDYVESQKRRSCCSF 222

Db 183 AGAGGPVKITENRKNKTSF 201

RESULT 7

US-09-822-860-5

; Sequence 5, Application US/09822860

; Patent No. US20020146795A1

; GENERAL INFORMATION:

; APPLICANT: ZHU, Shaoqing et al.

; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE

; TITLE OF INVENTION: PROTEINS, AND USES THEREOF

; FILE REFERENCE: CL001214

; CURRENT APPLICATION NUMBER: US/09/822,860

; CURRENT FILING DATE: 2001-04-02

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5

; LENGTH: 190

; TYPE: PRT

; ORGANISM: Discopyge ommata

US-09-822-860-5

Query Match 43.5%; Score 500; DB 10; Length 190;

Best Local Similarity 50.3%; Pred. No. 1.5e-46;

Matches 96; Conservative 42; Mismatches 51; Indels 2; Gaps 2;

Qy 27 DLTKGVMLLGDGTGKTCFLIOFKDGAFLSGTFTATVGIDFRNKVTVVDGVRVKLIQIWD 86

Db 1 DYLFKLLIGDSGVKTCCLFRSEDAF-NTTISTIGIDFKIRTVELDGKKIKLQIWD 59

Qy 87 AQGERFSVTHAYRDAQALLLLYDITNKSSFDNIRAWLTIHEYAQRDVVIMLGNKAD 146

Db 60 AQGERFRTITAYRGAMGIMKYDITNEKSFNKNWIRNIEEHASDVVERMILGNKCD 119

Qy 147 MSSERVIRSEDETLAREYGVPFLETSAKTGMNVELAFALAKELKYRAGHQADPSFOIR 205

Db 120 MNEKRVQSKERGEKLAIDYIGIKFLETSAKSSANVEEAFFTLARDIMTKLNKKNENSLQE 179

Qy 206 IRDYVESQKKR 216

Db 180 AVDKLKSPPKK 190

RESULT 8

US-09-925-300-1571

; Sequence 1571, Application US/09925300

; Patent No. US20020151681A1

; GENERAL INFORMATION:

; APPLICANT: Craig Rosen,

; APPLICANT: Steve Ruben

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA101

; CURRENT APPLICATION NUMBER: US/09/925,300

; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05988

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270


```

; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0367 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 57006
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-967-736-8

Query Match 41.0%; Score 472; DB 10; Length 201;
Best Local Similarity 43.1%; Pred. No. 1.7e-43;
Matches 87; Conservative 51; Mismatches 58; Indels 6; Gaps 3;

QY 23 SPSYDLTKVMMLGDTGKTCFLIQKDGAFSLGCTFIATVGDIFRNKVVTVVDGVRVKIQ 82
Db 2 NPEYDLFKLLIGDSGVGKSCLLLRFAADTY-TESYISTIGVDKIRTIELDGKTIKIQ 60

QY 83 IWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSFFDNIRAWLTIHEHYAQRDVVIMLIG 142
Db 61 IWDTAGQERFRVTSSYRGAGHIIIVYDVTDOESYANVKQLQIEDRYASENVNKLIVG 120

QY 143 NKADSSSERVIRSEGETLAREYGVFPFLETSAKTGMNVELAFIAIAKELKYRAGHOA--- 199
Db 121 NKSDLTKKVVNDTTAKFADSLGVPFLETSAKNATNVEQAPMTMAAEIKKRMGPGAASG 180

QY 200 -DEPSFQIRDYVESQKKRSSCC 220
Db 181 GERPNLKI-DSTPVKASGGCC 201

RESULT 12
US-09-967-736-3
; Sequence 3, Application US/09967736
; Patent No. US20020103340A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; Lal, preeti
; Corley, Neil C.
; Shah, Purvi
; TITLE OF INVENTION: RAB PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/967,736
; FILING DATE: 28-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/154,602
; FILING DATE: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0367 US
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LivTUT04
; CLONE: 2514506
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-967-736-3

Query Match 40.7%; Score 468; DB 10; Length 201;
Best Local Similarity 42.1%; Pred. No. 4.6e-43;
Matches 85; Conservative 53; Mismatches 58; Indels 6; Gaps 3;

QY 23 SPSYDLTKVMMLGDTGKTCFLIQKDGAFSLGCTFIATVGDIFRNKVVTVVDGVRVKIQ 82
Db 2 NPEYDLFKLLIGDSGVGKSCLLLRFAADTY-TESYISTIGVDKIRTIELDGKTIKIQ 60

QY 83 IWDTAGQERFRSVTHAYYRDAQALLLLYDITNKSFFDNIRAWLTIHEHYAQRDVVIMLIG 142
Db 61 IWDTAGQERFRVTSSYRGAGHIIIVYDVTDOESYANVKQLQIEDRYASENVNKLIVG 120

QY 143 NKADSSSERVIRSEGETLAREYGVFPFLETSAKTGMNVELAFIAIAKELKYRAGHOA--- 199
Db 121 NKSDLTKKVVNDTTAKFADSLGVPFLETSAKNATNVEQAPMTMAAEIKKRMGPGAASG 180

QY 200 -DEPSFQIRDYVESQKKRSSCC 220
Db 181 GERPNLKI-DSTPVKASGGCC 201

RESULT 13
US-09-794-257-16
; Sequence 16, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1e1
; FILE OF INVENTION: Human G-Proteins
; FILE REFERENCE: 35800/209285
; CURRENT APPLICATION NUMBER: US/09/794,257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam accession number PF00071
US-09-794-257-16

Query Match 40.5%; Score 466; DB 10; Length 198;
Best Local Similarity 47.1%; Pred. No. 7.4e-43;
Matches 97; Conservative 38; Mismatches 45; Indels 26; Gaps 6;

QY 31 KVMMLGDTGKTCFLIQKDGAFSLGCTFIATVGDIFRNKVVTVVDGVRVKIQIWDTAGOE 90
Db 1 KLVLLIGDSGVGKSCLLLRFDNKFVE-EYIPTIGVDFTYKTVVDGKTVKLQIWDTAGOE 59

QY 91 RFRSVTHAYYRDAQALLLLYDITNKSFFDNIRAWLTIHEHYAQRDVVIMLIGNKADMS 148
Db 60 RFRALPAYRCAQGFLLVYDITSRDSFENKVKLEILRHADKDNENVPLVGNKCDLE 119

QY 149 SE-----RVIRSEGETLAREYG-VPFLETSAKTGMNVELAFIAIAKELKYRAGH 197
Db 120 DDEDLLELEGQKRVVSTEBGEALAKELGALPEMETSAKTNTNVEEAFEELKVKVSE 179

```

QY 198 ---QADEPSFQIRDYVESOKKRSSCC 220
Db 180 VVNVLDQP-----AKKKSKCC 196

RESULT 14

US-09-945-173-5
; Sequence 5, Application US/09945173
; Patent No. US20020127568A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 47324, A NOVEL HUMAN G-PROTEIN AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 38155-20035.00
; CURRENT APPLICATION NUMBER: US/09/945,173
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,293
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-945-173-5

Query Match 40.5%; Score 466; DB 10; Length 198;
Best Local Similarity 47.1%; Pred. No. 7.4e-43;
Matches 97; Conservative 38; Mismatches 45; Indels 26; Gaps 6;

QY 31 KVMLLGDTGVGKTCFLIQFDGAFSLSGTFTATVGIDFRNKKVVTVDGVRVKLQIWDTAGQE 90
Db 1 KLVLLGDSGVGKSSLLIRFTDNKFVE-EYIPTIGVDYFTKTVEVDGKTVKLVQIWDTAGQE 59
QY 91 RFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLFEIHEYAQRD--VVIMLLGNKADMS 148
Db 60 RFRALRPAYYRGAQGFLLVDITSRDSFENVKKWLEELRHADKDNVPIVLVGNKCDLE 119
QY 149 SE-----RVIRSEDEGTALREYGVVPFLETSAKTGMNVVELAFLAIKELKYRAGH 197
Db 120 DDEDLLELTGQKRVVSTEEGALAKELGALPFMETSAKNTNVVEAFEELAREILKKVSE 179
QY 198 ---QADEPSFQIRDYVESOKKRSSCC 220
Db 180 VVNVLDQP-----AKKKSKCC 196

RESULT 15

US-09-972-529-4
; Sequence 4, Application US/09972529
; Patent No. US20020150916A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 47316, A NOVEL HUMAN G-PROTEIN AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 38155-20041.00
; CURRENT APPLICATION NUMBER: US/09/972,529
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 60/237,716
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence

US-09-972-529-4

Query Match 40.5%; Score 466; DB 10; Length 198;
Best Local Similarity 47.1%; Pred. No. 7.4e-43;
Matches 97; Conservative 38; Mismatches 45; Indels 26; Gaps 6;

QY 31 KVMLLGDTGVGKTCFLIQFDGAFSLSGTFTATVGIDFRNKKVVTVDGVRVKLQIWDTAGQE 90
Db 1 KLVLLGDSGVGKSSLLIRFTDNKFVE-EYIPTIGVDYFTKTVEVDGKTVKLVQIWDTAGQE 59
QY 91 RFRSVTHAYYRDAQALLLYDITNKSSFDNIRAWLFEIHEYAQRD--VVIMLLGNKADMS 148
Db 60 RFRALRPAYYRGAQGFLLVDITSRDSFENVKKWLEELRHADKDNVPIVLVGNKCDLE 119
QY 149 SE-----RVIRSEDEGTALREYGVVPFLETSAKTGMNVVELAFLAIKELKYRAGH 197
Db 120 DDEDLLELTGQKRVVSTEEGALAKELGALPFMETSAKNTNVVEAFEELAREILKKVSE 179
QY 198 ---QADEPSFQIRDYVESOKKRSSCC 220
Db 180 VVNVLDQP-----AKKKSKCC 196

Search completed: November 17, 2002, 16:43:36
Job time : 21 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 17, 2002, 13:58:08 ; Search time 2433 Seconds
(without alignments)
17799.725 Million cell updates/sec

Title: US-09-817-199A-1

Perfect score: 2674

Sequence: 1 ttccctgcggcgccgact.....aaaaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_est1.*

10: gb_est2.*

11: gb_htc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: gb_gss.*

18: em_gss_hum.*

19: em_gss_inv.*

20: em_gss_pln.*

21: em_gss_vrt.*

22: em_gss_fun.*

23: em_gss_mam.*

24: em_gss_mus.*

25: em_gss_Other.*

26: em_gss_pro.*

27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	921.6	34.5	1014	BM921365	BM921365 AGENCOURT
2	884.2	33.1	887	AL522282	AL522282 AL522282
C 3	847.4	31.7	855	AL580781	AL580781 AL580781
C 4	802.6	30.0	833	AL522281	AL522281 AL522281
5	755.4	28.2	881	BG759655	BG759655 602713348
6	741.6	27.7	868	BQ690583	BQ690583 AGENCOURT

7	733.2	27.4	761	13	BI819064
C 8	702	26.3	702	14	BM981939
9	701.2	26.2	1051	12	BG283602
C 10	696.8	26.1	700	14	BQ028124
11	643.6	24.1	1069	12	BG282782
C 12	626.4	23.4	629	14	BQ183276
13	605.4	22.6	607	9	AF188522
14	601.6	22.5	740	13	BI767046
15	600.6	22.5	691	12	BG253976
16	585.6	21.9	749	9	AL559085
17	581.2	21.7	886	14	BQ686932
18	553.2	20.7	676	10	BB598938
19	539.2	20.2	603	10	AW955318
20	529	19.8	529	14	BM706070
21	527.6	19.7	651	10	BB633978
22	516.8	19.3	655	10	BB635649
23	514.6	19.2	522	13	BM151643
C 24	503.4	18.8	507	13	BM151381
25	473.6	17.7	480	10	AW270880
26	464.4	17.4	493	9	AA324998
27	460.6	17.2	488	13	BM149118
28	442.4	16.5	676	10	BB206788
29	440.4	16.5	766	14	BQ188402
30	423.8	15.8	553	12	BE755280
31	422.4	15.8	733	10	BB619625
32	415.2	15.5	437	14	W36293
33	412.4	15.4	793	12	BG256040
C 34	409.4	15.3	434	9	AA806096
C 35	399.4	14.9	401	10	AW768344
36	397.4	14.9	657	10	BB196489
C 37	385	14.4	401	9	AA631787
C 38	381	14.2	403	9	AA878511
39	380.4	14.2	455	10	AA464449
40	377.8	14.1	385	9	AI243836
41	374.2	14.0	380	10	BE241772
42	362.8	13.6	579	10	BB621927
43	358	13.4	368	13	BM152052
44	347	13.0	348	9	AA310682
45	345.6	12.9	404	13	BM149006

ALIGNMENTS

RESULT 1
BM921365
LOCUS
DEFINITION
AGENCOURT_6626159 NIH_MGC_115 Homo sapiens cdna clone IMAGE:5752779
5', mRNA sequence.
ACCESSION
BM921365
VERSION
BM921365.1 GI:19371744
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1014)
NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12787 row: i column: 04
High quality sequence stop: 735.
Location/Qualifiers
1..1014
source

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:575279"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/notes="Organ: pooled brain, lung, testis; Vector:
pcmv-sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT      237 a 288 c 295 g 194 t
ORIGIN

Query Match      34.5%; Score 921.6; DB 14; Length 1014;
Best Local Similarity 96.7%; Pred. No. 5.2e-93;
Matches 973; Conservative 0; Mismatches 29; Indels 4; Gaps 3;

QY  2  TGCCTGCGGGCGGCACCTGCTCACTCTCGTCCAGGACATGACGGGACGCCAGCGCG 61
Db  8  TCGCCCTGCGGGCGGCACCTGCTCACTCTCGTCCAGGACATGACGGGACGCCAGCGCG 67
QY  62  CGTTGCCACCGGGATGGGAGGCGCCCGGAGGCTCCCGCCCTGCAGTCCGAGCTACGA 121
Db  68  CGTTGCCACCGGGATGGGAGGCGCCCGGAGGCTCCCGCCCTGCAGTCCGAGCTACGA 127
QY  122  CTTACGGGCAAGGTGATGCTTCTGGGAGACACAGCGCTGGGCAAAACATGTTCTCTGAT 181
Db  128  CTTACGGGCAAGGTGATGCTTCTGGGAGACACAGCGCTGGGCAAAACATGTTCTCTGAT 187
QY  182  CCAATTCGAAGAGGGGCGCTCTCTGTCGGAACCTTCATAGCCACCGCTGGCATAGACTT 241
Db  188  CCAATTCGAAGAGGGGCGCTCTCTGTCGGAACCTTCATAGCCACCGCTGGCATAGACTT 247
QY  242  CAGGAACAAGGTGATGCTGTGATGGCTGAGAGTGAAGCTGCAGATCTGGGACACCGC 301
Db  248  CAGGAACAAGGTGATGCTGTGATGGCTGAGAGTGAAGCTGCAGATCTGGGACACCGC 307
QY  302  TGGGCAAGAGCGTTCCGAAGCGTCAACCATGTTATACAGAGATGCTCAGGCGTTGCT 361
Db  308  TGGGCAAGAGCGTTCCGAAGCGTCAACCATGTTATACAGAGATGCTCAGGCGTTGCT 367
QY  362  TCTGCTGTATGACATCACCAAAATCTTCTTCGACAACATCAGGCGCTGGCTCACTGA 421
Db  368  TCTGCTGTATGACATCACCAAAATCTTCTTCGACAACATCAGGCGCTGGCTCACTGA 427
QY  422  GATTCATGATATGCCAGAGGACGTGGTGATCATGCTGTAGGCAACAAGGCGGATAT 481
Db  428  GATTCATGATATGCCAGAGGACGTGGTGATCATGCTGTAGGCAACAAGGCGGATAT 487
QY  482  GAGCAGGAAAGAGTATCGCTTCCGAAGCGGAGAGACCTTGGCCAGGAGTACGGTGT 541
Db  488  GAGCAGGAAAGAGTATCGCTTCCGAAGCGGAGAGACCTTGGCCAGGAGTACGGTGT 547
QY  542  TCCCTTCTCTGGAGACCGCGCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCAT 601
Db  548  TCCCTTCTCTGGAGACCGCGCAAGACTGGCATGAATGTGGAGTTAGCCTTTCTGGCCAT 607
QY  602  CGCAAGGAACTGAATATCCGGGCGGGGATCAGGCGGATGAGCCAGCTTCCAGATCCG 661
Db  608  CGCAAGGAACTGAATATCCGGGCGGGGATCAGGCGGATGAGCCAGCTTCCAGATCCG 667
QY  662  AGACTATGATAGTCCAGAAAGCGCTCCAGCTGCTCTCTCATGTGAATCCAGG 721
Db  668  AGACTATGATAGTCCAGAAAGCGCTCCAGCTGCTCTCTCATGTGAATCCAGG 727
QY  722  GGGCAGAGAGGAGGCTCTGAGGACACAGAGTGCAGGCTTCCCGCTCCAGGCGCTGGCT 781
Db  728  GGGCAGAGAGGAGGCTCTGAGGACACAGAGTGCAGGCTTCCCGCTCCAGGCGCTGGCT 787
```

```
QY  782  TATTCCAAGAGGCTGAGCCAAT-GGGGAGAAAGATGGAGACTACTGTCACAGCCGCTTC 840
Db  788  TATTCCAAGAGGCTGAGCCAATGGGGAGAAAGATGGAGACTACTGTCACAGCCGCTTC 847
QY  841  CTAGCAGGAGCTATACTCCAACTCTTACTTGATGTTCTTCGGTCTCCCGCAT--CCAC 898
Db  848  CTAGCAGGAGCTATACTCCAACTCTTACTTGATGTTCTTCGGTCTCCCGCATTCACCC 907
QY  899  AGGGAGGCTAAACACTTACCTTATTTTAAATAGTACATAATTTAATACCAAAAAA-GG 957
Db  908  GGAAGGCTAAACACTTACCTTATTTTAAATAGTACATAATTTAATACCAAAAAAAGG 967
QY  958  CGCCTGGATCCCAAAAAACCGAGGCTGGAGCTAGTGGCCCTTTT 1003
Db  968  CGCCTGGATCCCAAAAAACCGAGGCTGGAGCTAGTGGCCCTTT 1013

RESULT 2
AL522282
LOCUS      AL522282 LTI_NFL004_NBC2 887 bp mRNA linear EST 13-FEB-2001
DEFINITION prime, mRNA sequence.
ACCESSION AL522282
VERSION   AL522282.1 GI:12785775
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 887)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES             source
     1..887
         /organism="Homo sapiens"
         /db_xref="taxon:9606"
         /clone="CS0DB008YF08"
         /clone_lib="LTI_NFL004_NBC2"
         /sex="male"
         /tissue_type="neuroblastoma cells"
         /lab_host="DH10B"
         /note="Organ: brain; Vector: PCMVSPORT 6; 1st strand cDNA
         was primed with a NotI-oligo(dT) primer. Five prime end
         enriched, double-stranded cDNA was digested with Not I and
         cloned into the Not I and Eco RV sites of the PCMVSPORT 6
         vector. Library was normalized. Library was constructed
         by Life Technologies. Contact : Feng Liang Life
         Technologies, a division of Invitrogen 9800 Medical Center
         Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
         8371 Email : fliang@lifetech.com URL :
         http://fulllength.invitrogen.com"

BASE COUNT      203 a 256 c 257 g 168 t 3 others
ORIGIN

Query Match      33.1%; Score 884.2; DB 9; Length 887;
Best Local Similarity 99.5%; Pred. No. 7.4e-89;
Matches 883; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY  15  GGCACCTGCTCACTCTCGTCCAGGACATGACGGGACGCGCAGCGCGGTGCCACCCGG 74
Db  1  GGCACCTGCTCACTCTCTCTCCAGGACATGACGGGACGCGCAGCGCGGTGCCACCCGG 60
QY  75  GATGGCAGGCGCCCGGAGGCGCTCCCGCCCTGCAGTCCGAGCTACGACCTACAGGGCAAG 134
Db  61  GATGGCAGGCGCCCGGAGGCGCTCCCGCCCTGCAGTCCGAGCTACGACCTACAGGGCAAG 120
QY  135  GTGATCCTCTCTGGGAGACACAGCGCTCGGCAAAACATGTTCTCTGATCAATTCAGAC 194
```

```
|||||
Db 121 GTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCCCTGATCCAATTCAAAGAC 180
Qy 195 GGGGCGCTTCCTGTCGCGAACCCTTCATAGCCACCCCTCGCATAGACTTCAGGAACAAGGTG 254
Db 181 GGGGCGCTTCCTGTCGCGAACCCTTCATAGCCACCCTCGCATAGACTTCAGGAACAAGGTG 240
Qy 255 GTGACTGTGGATGCGGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGGCGAGCAAGCG 314
Db 241 GTGACTGTGGATGCGGTGAAGTGAAGCTGCAGATCTGGGACACCGCTGGGCGAGCAAGCG 300
Qy 315 TTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCGCTTCTCTCTCTCTATGAC 374
Db 301 TTCCGAAGCGTCACCCATGCTTATTACAGARATGCTCAGGCGCTTCTCTCTCTCTATGAC 360
Qy 375 ATCACCAACAATCTCTTCGCAACAATCAGGCGCTGGCTCAGTGCAGATTCATGAGTAT 434
Db 361 ATCACCAACAATCTCTTCGCAACAATCAGGCGCTGGCTCAGTGCAGATTCATGAGTAT 420
Qy 435 GCCCAGAGGACGCTGTGATCATGCTGTAGGCAACAGCGGATATGAGCAGGCAAGA 494
Db 421 GCCCAGAGGACGCTGTGATCATGCTGTAGGCAACAGCGGATATGAGCAGGCAAGA 480
Qy 495 GTGATCGTTCGGAAGAGGAGACACCTTGGCCAGGAGTACGGTGTTCCTTCCCTGGAG 554
Db 481 GTGATCGTTCGGAAGAGGAGACACCTTGGCCAGGAGTACGGTGTTCCTTCCCTGGAG 540
Qy 555 ACCAGCGCGGCGGATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAG 614
Db 541 ACCAGCGCGGCGGATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAG 600
Qy 615 AAATACCGGCGGCGGATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAG 674
Db 601 AAATACCGGCGGCGGATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAG 660
Qy 675 TCCCAGAGAACGCTCAGCTGCTGCTCTTCATGTGAATCCAGGGGGCAGAGAGAG 734
Db 661 TCCCAGAGAACGCTCAGCTGCTGCTCTTCATGTGAATCCAGGGGGCAGAGAGAG 720
Qy 735 GCTCTGAGGACACAGACTCAGCGCTTCCCGCTCCAGCGCTGGCTTATTCAGAGGC 794
Db 721 GCTCTGAGGACACAGACTCAGCGCTTCCCGCTCCAGCGCTGGCTTATTCAGAGGC 780
Qy 795 TGAGCCAAATGGGGAAGAGTGGAGGACTCACTCACAGCGCGCTTCCCTAGCAGGAGCTA 854
Db 781 TGAGCCAAATGGGGAAGAGTGGAGGACTCACTCACAGCGCGCTTCCCTAGCAGGAGCTA 840
Qy 855 TACTCCAACTCTACTTGAGTTCCTGCGGTCTCCCGGATCCACAGG 901
Db 841 TACTCCAACTCTACTTGAGTTCCTGCGGTCTCCCGGATCCACAGG 887

RESULT 3
AL580781/c AL580781 855 bp mRNA linear EST 16-FEB-2001
LOCUS AL580781 LTI_NFL008_Tc2 Homo sapiens cdna clone CS0DJ014YB11 3
DEFINITION prime, mRNA sequence.
ACCESSION AL580781
VERSION AL580781.1 GI:12947137
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 855)
LI,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cdna libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
```

```
source
1. .855
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DJ014YB11"
/clone_lib="LTI_NFL008_Tc2"
/sex="male"
/tissue_type="T cells from T cell leukemia"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with NotI and
cloned into the NotI and EcoRV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies, Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 204 a 224 c 224 g 201 t 2 others
ORIGIN
Query Match 31.7%; Score 847.4; DB 9; Length 855;
Best Local Similarity 99.68; Pred. No. 8.5e-85;
Matches 848; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1700 ATTTTTCGTAGACACAGGGTTTCGCCATGTTGCCAGGCTGGTCTTGAATTCCTGAGCTC 1759
Db 855 ATTTTTCGTAGACACAGGGTTTCGCCATGTTGCCAGGCTGGTCTTGAATTCCTGAGCTC 796
Qy 1760 AAGCAACCTGCGGCGCTCGGCCATCCCAAAGTACTGGGATTTACACGACAGGACCATGTC 1819
Db 795 AAGCAACCTGCGGCGCTCGGCCATCCCAAAGTACTGGGATTTACACGACAGGACCATGTC 736
Qy 1820 CCAGCTAGATGTGCTTATCCCAATCTTTGGCAGGATGCGCTCCACAGGCGATTTTC 1879
Db 735 CCAGCTAGATGTGCTTATCCCAATCTTTGGCAGGATGCGCTCCACAGGCGATTTTC 676
Qy 1880 TTCAGCAGCTGAAGTGTTCAGCCCTCTGGGTTAAGAGCCAGATAGAGAGAAATCCCTT 1939
Db 675 TTCAGCAGCTGAAGTGTTCAGCCCTCTGGGTTAAGAGCCAGATAGAGAGAAATCCCTT 616
Qy 1940 TCCTAGGTTTGAATGTGTTGTAAGAAAAGAGAAATCCCTGGCTCTCTGGAGCTGGTG 1999
Db 615 TCCTAGGTTTGAATGTGTTGTAAGAAAAGAGAAATCCCTGGCTCTCTGGAGCTGGTG 556
Qy 2000 GAGACAAGATTAAAGCAAACTCCCTCGATGTATCCCTTTGACCCCAAGCTCTGCCTCC 2059
Db 555 GAGACAAGATTAAAGCAAACTCCCTCGATGTATCCCTTTGACCCCAAGCTCTGCCTCC 496
Qy 2060 TCCCTGACACCCATGCGCTTTTCCTTTAACTTCTCAAAACAGATACAGGCGCTAAACTGC 2119
Db 495 TCCCTGACACCCATGCGCTTTTCCTTTAACTTCTCAAAACAGATACAGGCGCTAAACTGC 436
Qy 2120 TTTACCTCCCTCTCTAGTGCAGTTAGTGGTGGAGGTCACCCATTTCCGAGTTAA 2179
Db 435 TTTACCTCCCTCTCTAGTGCAGTTAGTGGTGGAGGTCACCCATTTCCGAGTTAA 376
Qy 2180 ACCAATGCAATATGATGTAACAAAGTCAATGCTGCGGTATGCTGCGGTAGAGAGGGGTA 2239
Db 375 ACCAATGCAATATGATGTAACAAAGTCAATGCTGCGGTATGCTGCGGTAGAGAGGGGTA 316
Qy 2240 GCAAGTTTCATGTCTCTCTTGGTCACATATCTCCAAAGCTCTGATCCCTGCCATGGGA 2299
Db 315 GCAAGTTTCATGTCTCTCTTGGTCACATATCTCCAAAGCTCTGATCCCTGCCATGGGA 256
Qy 2300 AGTGACAGGAACATGAGTGCATGACCTGACGAGGATCTTTACTCAGCTCTGCGGCGCT 2359
Db 255 AGTGACAGGAACATGAGTGCATGACCTGACGAGGATCTTTACTCAGTCTCTGCGGCGCT 196
Qy 2360 GGAGGGGAGAGGGGAGGAAGATGTCGCTGCACATTTCTGAGGCTTACTGCAATTTGC 2419
Db 195 GGAGGGGAGAGGGGAGGAAGATGTCGCTGCACATTTCTGAGGCTTACTGCAATTTGC 136
Qy 2420 TTTCAAGGCAGAAATCTTCTGCTGTAGCAGTCAAGCGGCTCCAGTTTGGGCCCCATAGGAA 2479
```

```
Db 135 TTTCAAGGAGAGAAATCTTCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCCCATAGGAA 76
|||||
Qy 2480 GTTCTCCGTGGCTCCTCAGGACAGCAGGAGGAGGAGGTGACATTTGCCAGTCTCTTCTG 2539
|||||
Db 75 GTTCTCCGTGGCTCCTCAGGACAGCAGGAGGAGGAGGTGACATTTGCCAGTCTCTTCTG 16
|||||
Qy 2540 GGGCCCAAGGC 2550
|||||
Db 15 GGGCCCAAGGC 5
|||||

RESULT 4
AL522281/c 833 bp mRNA linear EST 13-FEB-2001
LOCUS AL522281.LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB008YF08 3
DEFINITION prime, mRNA sequence.
ACCESSION AL522281
VERSION AL522281.1 GI:12785774
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 833)
Li.W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr.
Location/Qualifiers
1. 833
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CS0DB008YF08"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="PH108"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with NotI and
cloned into the NotI and EcoRV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact: Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax: (1) 301 610
8371 Email: filiang@lifetech.com URL:
http://fulllength.invitrogen.com"

BASE COUNT 206 a 207 c 215 g 200 t 5 others
ORIGIN

Query Match 30.0%; Score 802.6; DB 9; Length 833;
Best Local Similarity 98.7%; Pred. No. 7.3e-80;
Matches 814; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

Qy 1673 GTACATCACACCTGGCTAATTTTGTATTTTGTAGACACAGGGTTTCGCCATGTTCG 1732
|||||
Db 825 KTACCATCACACCTGGCTAATTTTGTATTTTGTAGACACAGGGTTTCGCCATGTTCG 766
|||||
Qy 1733 CCAGGCTGGTCTTTGAATTCCTGAGCTCAAGCAACCTGCCGGCTCGCCCTCCCAAGTAC 1792
|||||
Db 765 CCAGGCTGGTCTTTGAATTCCTGAGCTCAAGCAACCTGCCGCCCTCGCCYCCCAAGTAC 706
|||||
Qy 1793 TGGGATTACACGACAGGACCATGCCAGGCTAGCTGTCTTATCCCAATCCCTTGG 1852
|||||
Db 705 TGGGATTACACGACAGGACCATGCCAGGCTAGCTGTCTTATCCCAATCCCTTGG 646
|||||
Qy 1853 CAGCATGACAGCTCCACAGCGGATTTCTTCAAGCAGCTGAAGTGTATAGCCCTCCCTGGGT 1912
|||||
Db 645 CAGCATGACAGCTCCACAGCGGATTTCTTCAAGCAGCTGAAGTGTATAGCCCTCCCTGGGT 586
|||||
```

```
Qy 1913 TAAGAGCCAGATAAGAGAGAAATCCCTTTCTTAGGTTTGAATGTTGTGAAAAAAGA 1972
|||||
Db 585 TAAGAGCCAGATAAGAGAGAAATCCCTTTCTTAGGTTTGAATGTTGTGAAAAAAGA 526
|||||
Qy 1973 GAAATCCCTGGCTCCTGGAGCTGGTGGAGACAAGATTAAAGCAAAACCTCCCTCGACATGT 2032
|||||
Db 525 GAAATCCCTGGCTCCTGGAGCTGGTGGAGACAAGATTAAAGCAAAACCTCCCTCGACATGT 466
|||||
Qy 2033 ATCCCTTTGACCCCAAGCTCTGCCTCCCTCCCTGACACCCATGCCCTTTCCCTTAACTTC 2092
|||||
Db 465 ATCCCTTTGACCCCAAGCTCTGCCTCCCTCCCTGACACCCATGCCCTTTCCCTTAACTTC 406
|||||
Qy 2093 TCAAAACAGATACACAGGCGCTAAACTGCTTTACCTCCCTCCCTACTG-AGTCAAGTTAGT 2151
|||||
Db 405 TCAAAACAGATACACAGGCGCTAAACTGCTTTACCTCCCTCCCTACTG-AGTCAAGTTAGT 346
|||||
Qy 2152 GGTGGAGGTCAACCATTTCCCGAGTTAAACCAATCAATATAGTAAACAAGTCAATGT 2211
|||||
Db 345 GGTGGAGGTCAACCATTTCCCGAGTTAAACCAATCAATATAGTAAACAAGTCAATGT 286
|||||
Qy 2212 GGGTATGCTGGGTAGAGAGAGGGGTAGCAAGTTTCATGTGCTCCTCTGGTCAATATC 2271
|||||
Db 285 GGGTATGCTGGGTAGAGAGAGGGGTAGCAAGTTTCATGTGCTCCTCTGGTCAATATC 226
|||||
Qy 2272 TCCAAAAGCTCTGATCCCTGTCATGGGAAGTGGACAGGAAACATGAGTCACTGCA 2331
|||||
Db 225 TCCAAAAGCTCTGATCCCTGTCATGGGAAGTGGACAGGAAACATGAGTCACTGCA 166
|||||
Qy 2332 GCATCTTTTACTGCAGCTCTGCCGCTGGAGGGGAGAGGAGGAAGTATGCGC 2391
|||||
Db 165 GCATCTTTTACTGCAGCTCTGCCGCTGGAGGGGAGAGGAGGAAGTATGCGC 106
|||||
Qy 2392 TGCACATTTCTGAGGCTACTGCTATTTGCTTCAAGSCAGAAATCTTGCTCTGAGCAGTCA 2451
|||||
Db 105 TGCACATTTCTGAGGCTACTGCTATTTGCTTCAAGSCAGAAATCTTGCTCTGAGCAGTCA 46
|||||
Qy 2452 GGGGCTCCAGTTTGGGCCCGGATAAGGAAGTTCTCGGTGGCCCTCC 2496
|||||
Db 45 GCGGCTCCAGTTTGGGCCCGGATAAGGAAGTTCTCGGTGGCCCTCC 1
|||||

RESULT 5
Bg759655 881 bp mRNA linear EST 15-MAY-2001
LOCUS 602713348F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853626 5',
DEFINITION mRNA sequence.
ACCESSION Bg759655
VERSION Bg759655.1 GI:14070308
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 881)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC1699 row: p column: 11
High quality sequence stop: 707.
Location/Qualifiers
1. 881
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source
```



```
/clone="IMAGE:4853626"
/clone_lib="NIH_MGC_48"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/site="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT      220 a      236 c      261 g      164 t
ORIGIN

Query Match      28.2%; Score 755.4; DB 12; Length 881;
Best Local Similarity 94.6%; Pred. No. 1.1e-74;
Matches 805; Conservative 0; Mismatches 41; Indels 5; Gaps 2;

Qy 566 GACTGGCATGAATGTGAGTTAGCTTTCTGGCCATCGCCAAAGCACTGAAATACCGGGC 625
Db 2 GACTGGCATGAATGTGAGTTAGCTTTCTGGCCATCGCCAAAGCACTGAAATACCGGGC 61

Qy 626 CGGSCATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAA 685
Db 62 CGGSCATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAA 121

Qy 686 CGGTCAGCTGCTGCTTCATGTGTAATCCAGGGGCGAGAGAGGCTCTGGAGGC 745
Db 122 CGGTCAGCTGCTGCTTCATGTGTAATCCAGGGGCGAGAGAGGCTCTGGAGGC 181

Qy 746 ACACAGATGACGCTTCCCTCCAGGCTGCTTATCCAGAGGCTCAGCAATGG 805
Db 182 ACACAGATGACGCTTCCCTCCAGGCTGCTTATCCAGAGGCTCAGCAATGG 241

Qy 806 GGAGAAAGATGGAGGACTCACTGCACAGCGCTTCTTAGCAGGAGCTATACTCAACTC 865
Db 242 GGAGAAAGATGGAGGACTCACTGCACAGCGCTTCTTAGCAGGAGCTATACTCAACTC 301

Qy 866 CTACTTGAATGCTGCGGCTCCCGGATCCACAGGAGGGTAAACACATTAGCTTTTAT 925
Db 302 CTACTTGAATGCTGCGGCTCCCGGATCCACAGGAGGGTAAACACATTAGCTTTTAT 361

Qy 926 TTTAATAGTACATAATTTATACCAAAAGGCGCTTGATCCCAAAAGGAGGCTG 985
Db 362 TTTAATAGTACATAATTTATACCAAAAGGCGCTTGATCCCAAAAGGAGGCTG 421

Qy 986 GGAGCTAGTGGCCCTTTTGTCTTCTAGGACTTTGGGGGCGCGCCCTCCTCTAAGCATA 1045
Db 422 GGAGCTAGTGGCCCTTTTGTCTTCTAGGACTTTGGGGGCGCGCCCTCCTCTAAGCATA 481

Qy 1046 ACAAGGTGGTGTCTCCAGCTCAGCCCGCAGGGGACACAGATGCATTTGGGGGTGAGG 1105
Db 482 ACAAGGTGGTGTCTCCAGCTCAGCCCGCAGGGGACACAGATGCATTTGGGGGTGAGG 541

Qy 1106 GCAGGTAATGACTCCATCGACCTCAGTTCAGCTGCGACAGGCTCAGGTGACCCAGC 1165
Db 542 GCAGGTAATGACTCCATCGACCTCAGTTCAGCTGCGACAGGCTCAGGTGACCCAGC 601

Qy 1166 CTTTCACTGTCTCCGGCTCTCCAGGAGCTTATCTTCCGCCCATCTCCCAAAATAGTGGGCC 1225
Db 602 CTTTCACTGTCTCCGGCTCTCCAGGAGCTTATCTTCCGCCCATCTCCCAAAATAGTGGGCC 661

Qy 1226 CTTTGTGTGTGAGGAAGACCAAGCCCTCAGGGAAGATATGAGATGGAGGGG 1285
Db 662 CTTTGTGTGTGAGGAAGACCAAGCCCTCAGGGAAGATATGAGATGGAGGGG 721

Qy 1286 GAGCACAAAGGGGAGAGTAGGTCTA ---GCTGGCTATCTCTGGCCCTTACTAACACC 1342
Db 722 GAGCAAAGGGGAGAGTAGGTCTAAGGCAAGCATGCGCAAGTCTCTGGCCGTACAAGAGCG 781

Qy 1343 CCCTGG --AGSCATGCCCCCTTTTCTCCAGCACACAAGCACATTGGGGCACCTTGGAAATAT 1400
```

```
Db 782 CCCTGGGAGGCGATGCCCTTATAAGCCAGGACACAGGCCATTTGGGCACCTTGGAAATAG 841
Qy 1401 TGGTTCAGGC 1411
Db 842 TGGACCGGAGC 852

RESULT 6
LOCUS      B0690583      868 bp      mRNA      linear      EST 15-JUL-2002
DEFINITION AGENCOURT_8186311 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6251642
5', mRNA sequence.
ACCESSION  B0690583
VERSION    B0690583.1 GI:21815899
KEYWORDS   EST.
SOURCE      Human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 868)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: rcgaps@nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LCM2396 row: k column: 03
            High quality sequence stop: 532.

FEATURES             Location/Qualifiers
     source            1..868
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6251642"
                     /clone_lib="NIH_MGC_110"
                     /tissue_type="ductal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
BASE COUNT      215 a      239 c      240 g      172 t      2 others
ORIGIN

Query Match      27.7%; Score 741.6; DB 14; Length 868;
Best Local Similarity 93.6%; Pred. No. 3.5e-73;
Matches 805; Conservative 0; Mismatches 51; Indels 4; Gaps 3;

Qy 490 AAAGAGTGCATCGTTCCGAAGCAGGAGACCTTGGCCAGGAGTACGGTGTTCCTTCC 549
Db 1 AAAGAGTGCATCGTTCCGAAGCAGGAGACCTTGGCCAGGAGTACGGTGTTCCTTCC 60

Qy 550 TGGAGACCGCCCAAGACTGGCATGAATGTGGAGTTAGCCCTTCTGCGCATCGCAAGG 609
Db 61 TGGAGACCGCCCAAGACTGGCATGAATGTGGAGTTAGCCCTTCTGCGCATCGCAAGG 120

Qy 610 AACTGAATATACCGCGCGGATCATAGCGGATGATGCCAGCTTCCAGATCCCGAGACTATG 669
Db 121 AACTGAATATACCGCGCGGATCATAGCGGATGATGCCAGCTTCCAGATCCCGAGACTATG 180

Qy 670 TAGAGTCCCAAGAGCGCTCCAGCTGCTGCTCTCTCATGTGAATCCAGGGGCGCAGAG 729
Db 181 TAGAGTCCCAAGAGCGCTCCAGCTGCTGCTCTCTCATGTGAATCCAGGGGCGCAGAG 240
```

QY	730	AGGAGGCTGTGAGGCACACAGGATGCAGCCTTCCCTCCAGGCTGGCTTATTC	789
Db	241	AGGAGGCTGTGAGGCACACAGGATGCAGCCTTCCCTCCAGGCTGGCTTATTC	300
QY	790	GAGGCTGAGCAATGGGAGAAAGATGAGGACTCACTGCACAGCGCTTCTAGCAGG	849
Db	301	GAGGCTGAGCAATGGGAGAAAGATGAGGACTCACTGCACAGCGCTTCTAGCAGG	360
QY	850	AGCTATACCAACCTCTACTGAGTCTCGGCTCCCTCCATCCACAGGAGGTAA	909
Db	361	AGCTATACCAACCTCTACTGAGTCTCGGCTCCCTCCATCCACAGGAGGTAA	420
QY	910	AACACTTAGCTTTTATTTAATAGTACATAATTTAATACAAAAGGCGCTGGATCC	969
Db	421	AACACTTAGCTTTTATTTAATAGTACATAATTTAATACAAAAGGCGCTGGATCC	480
QY	970	CAAAAACCGAGCTGGGAGCTAGTGGCCCTTTTGTCTTAGGACTT-GGGGGGCGGC	1028
Db	481	CAAAAACCGAGCTGGGAGCTAGTGGCCCTTTTGTCTTAGGACTT-GGGGGGCGGC	540
QY	1029	CTCTCCCTCTAAGCATAAACAAAGGTGGTGTCTCCAGCTCAGCCCGGACACAGAT	1088
Db	541	CTCTCCCTCTAAGCATACNAAGTGGTGTCTCCAGCTCAGCCCGGACACAGAT	600
QY	1089	GCACCTTT-GGGGGTGAGGCGAGTAAATGACTCCATCGCACCTCAGTTCAGTGACAGA	1147
Db	601	GCACCTTTGGGGGTGAGGCGAGTAAATGACTCCATCGCACCTCAGTTCAGTGACAGA	660
QY	1148	GGCTAGGTGACCCAGCCTTCACTGTCTCCGCTCTCCAGAGCTTATCTCGCCCAT	1207
Db	661	GGCTAGGTGACCCAGCCTTCACTGTCTCCGCTCTCCAGAGCTTATCTTCGCCCAT	720
QY	1208	CTCCCAATAAGTGGGCCCTTGTGTGTGAGGAAGCAAGCCCTCAGGGAAGATAAGAG	1267
Db	721	CTCCCAATAAGTGGGCCCTTGTGTGTGAGGAAGCAAGCCCTCAGGGAATAAGAA	780
QY	1268	ATATGAGATGAGGAGGAGGACAAAGGG--CAGAGATAGGGTCTAGCTGGCTATCTC	1325
Db	781	TATGAAAGTGAAGGGGGAAGCAAGGGGGCAAAAAGAGGCTTAACCTGGCTTATC	840
QY	1326	TGGCCTTACTACACCCCC 1345	
Db	841	TCTTGGCCTTAACCTACACCC 860	
RESULT	7		
LOCUS	BI819064		
DEFINITION	60303417f1 NIH_MGC_115 Homo sapiens cdna clone IMAGE:5174308 5', mRNA linear EST 04-OCT-2001		
ACCESSION	BI819064		
VERSION	BI819064.1		GI:15930614
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 761)		
TITLE	NIH-MGC http://mgc.ncl.nih.gov/		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing By: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM11434 row: b column: 05 High quality sequence stop: 759.		

FEATURES	Location/Qualifiers
source	1. .761
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="IMAGE:5174308"
	/clone_lib="NIH_MGC_115"
	/lab_host="DH10B"
	/note="Organ: pooled brain, lung, testis; Vector: pCMV-Sport6; Site: 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."
BASE COUNT	174 a 219 c 204 g 164 t
ORIGIN	
	Query Match 27.4%; Score 733.2; DB 13; Length 761;
	Best Local Similarity 99.3%; Pred. No. 3.3e-72;
	Matches 757; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
QY	728 AGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCTCCAGGCTGGCTTATTC 787
Db	1 AGAGGAGGCTCTGGAGGCACACAGGATGCAGCCTTCCCTCCAGGCTGGCTTATTC 60
QY	788 AAGAGGCTGAGCAATG-CGGAGAAAGATGGAGGACTACTGCACAGCGCTTCCTAGCA 846
Db	61 AAGAGGCTGAGCAATGTTGGGAGAAAGATGGAGGACTACTGCACAGCGCTTCCTAGCA 120
QY	847 GGGAGCTATACCTCAACTCTACTTCTCTGCGGTCTCCCGCATCCACAGGAGGG 906
Db	121 GGGAGCTATACCTCAACTCTACTTCTCTGCGGTCTCCCGCATCCACAGGAGGG 180
QY	907 TAAACACTTAGCTTTTATTTAATAGTACATAATTTAATACAAAGAGGCGCTGGAT 966
Db	181 TAAACACTTAGCTTTTATTTAATAGTACATAATTTAATACAAAGAGGCGCTGGAT 240
QY	967 CCCCCAAAACCGAGCTGGGAGCTAGTGCCCTTTTGTCTTCTAGGACTTGGGGGGCG 1026
Db	241 CCCCCAAAACCGAGCTGGGAGCTAGTGCCCTTTTGTCTTCTAGGACTTGGGGGGCCA 300
QY	1027 GCCCTCCCTCTTAAGCATAAACAAAGTGGTGTGTCTCCAGCTCAGCCCGGAGACACAG 1086
Db	301 GCCCTCCCTCTTAAGCATAAACAAAGTGGTGTGTCTCCAGCTCAGCCCGGAGACACAG 360
QY	1087 ATGCACCTTTGGGGTGGAGGCAAGTAAATGACTCCATCGCACCTCAGTTCAGCTGACAG 1146
Db	361 ATGCACCTTTGGGGTGGAGGCAAGTAAATGACTCCATCGCACCTCAGTTCAGCTGACAG 420
QY	1147 AGGCTCAGGTGACCCAGCCTTCTACTGTCTCCGCTCTCCAGGAGCTTATCTTCGCCCA 1206
Db	421 AGGCTCAGGTGACCCAGCCTTCTACTGTCTCCGCTCTCCAGGAGCTTATCTTCGCCCA 480
QY	1207 TCTCCAAAATAAGTGGGCCCTTGTGTGTGAGGAAGACCAAGCCCTCAGGGAAGATAAGA 1266
Db	481 TCTCCAAAATAAGTGGGCCCTTGTGTGTGAGGAAGACCAAGCCCTCAGGGAAGATAAGA 540
QY	1267 GATATGGAGATGGAGGGGAGGACAAAGGGCAGAGTAGGGTCTAGCTGGCTATCTCT 1326
Db	541 GATATGGAGATGGAGGGGAGGACAAAGGGCAGAGTAGGGTCTAGCTGGCTATCTCT 600
QY	1327 GGCCTTACTAACACCCCTCGGAGGATCCCTTTTCTCCAGCACACAAGACATTTGG 1386
Db	601 GGCCTTACTAACACCCCTCGGAGGATCCCTTTTCTCCAGCACACAAGACATTTGG 660
QY	1387 GCACCTTGAATAATTTGGTTCAGGCTTCCTGTCTCTGGACTTCAGATCTCGGGGAGCCC 1446
Db	661 GCACCTTGAATAATTTGGTTCAGGCTTCCTGTCTCTGGACTTCAGATCTCGGGGAG-CC 719
QY	1447 CTCGCCCTCTGAATCCCTGGCTTAGCTACCTACCTTCCTGCTGT 1488

Db	720	CTCCCCCCTGAATCCCTGGCTTACCTACCTCTCGCTGT	761	
RESULT 8				
BM981939/c				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
COMMENT				
FEATURES				
source				
BASE COUNT				
ORIGIN				
Query Match				
Best Local Similarity				
Matches 702;				
Conservative				
Mismatches				
Indels				
Gaps				
0;				
1953				
ATGTGTTGTGAAAAAAGAGAAATCCCTGGCTCTCGCTGGAGTGGTGGGAGACAAGATTAA				
2012				
643				
702				
ATGTGTTGTGAAAAAAGAGAAATCCCTGGCTCTCGCTGGAGTGGTGGGAGACAAGATTAA				
2013				
GCAAACTCCCTCGACATGATATCCCTTTTGACCCCAAGCTCTGCTCCCTGACCAACC				
2072				
583				
642				
GCAAACTCCCTCGACATGATATCCCTTTTGACCCCAAGCTCTGCTCCCTGACCAACC				
2073				
ATGCCCTTTCCCTTTAACTTCTCAAAACAGATACACAGGGCTTAACTGCTTTACCTCCCTC				
2132				
582				
ATGCCCTTTCCCTTTAACTTCTCAAAACAGATACACAGGGCTTAACTGCTTTACCTCCCTC				
523				
2133				
CTACTGAGTCAGGTTAGGTTGGGAGGTACCCATTTCCGAGTTTAAACCAATGCAATAT				
2192				
522				
CTACTGAGTCAGGTTAGGTTGGGAGGTACCCATTTCCGAGTTTAAACCAATGCAATAT				
463				
2193				
GAGTAAACAAAGTCTATGCTGCTGGGTAGAGAGAGGGGTAGCAAGTTTCATGCTG				
2252				
462				
GAGTAAACAAAGTCTATGCTGCTGGGTAGAGAGAGGGGTAGCAAGTTTCATGCTG				
403				
2253				
TCCTCTTGGTCAATATCTCCAAAGCTCTGATCCCTGCCATGGAAAGTGACACAGGAAA				
2312				
402				
TCCTCTTGGTCAATATCTCCAAAGCTCTGATCCCTGCCATGGAAAGTGACACAGGAAA				
343				
2313				
CATGAGGTCTATGACCTGACAGGATCTTTACTGACGCTCTGCCGCTCGAGGGGAGAGG				
2372				
342				
CATGAGGTCTATGACCTGACAGGATCTTTACTGACGCTCTGCCGCTCGAGGGGAGAGG				
283				
2373				
GGGAGAAGAAGTATGCTGCTGCACATTTCTGAGGCTACTGCAATTTGCTTCAAGGCAGAA				
2432				
282				
GGGAGAAGAAGTATGCTGCTGCACATTTCTGAGGCTACTGCAATTTGCTTCAAGGCAGAA				
223				
2433				
ATCTTCTCTGACGACGCTGACGCTCCAGTTTGGGCGCGATTAAGGAAGTTCTCCGTGCC				
2492				
222				
ATCTTCTCTGACGACGCTGACGCTCCAGTTTGGGCGCGATTAAGGAAGTTCTCCGTGCC				
163				
2493				
TCCTCAGGACGACGAGGAGGAGGCTGACATTTGCGAGTCTCTCTGCGGGCCCAAGGCAG				
2552				
162				
TCCTCAGGACGACGAGGAGGAGGCTGACATTTGCGAGTCTCTCTGCGGGCCCAAGGCAG				
103				
2553				
GTTCAGGAGATCCATCCATAGACGCTCTGGGCGCTCTGCAATTTGCTTCAAGTTTCAGAA				
2612				
102				
GTTCAGGAGATCCATCCATAGACGCTCTGGGCGCTCTGCAATTTGCTTCAAGTTTCAGAA				
43				
2613				
TTAACTGCAAGTATTTTGGAAAGCAAAAAAAGAAAAA				
2654				
42				
TTAACTGCAAGTATTTTGGAAAGCAAAAAAAGAAAAA				
1				
RESULT 9				
BM981939				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				
BASE COUNT				
ORIGIN				
Query Match				
Best Local Similarity				
Matches 702;				
Conservative				
Mismatches				
Indels				
Gaps				
0;				
1953				
ATGTGTTGTGAAAAAAGAGAAATCCCTGGCTCTCGCTGGAGTGGTGGGAGACAAGATTAA				
2012				
643				
702				
ATGTGTTGTGAAAAAAGAGAAATCCCTGGCTCTCGCTGGAGTGGTGGGAGACAAGATTAA				
2013				
GCAAACTCCCTCGACATGATATCCCTTTTGACCCCAAGCTCTGCTCCCTGACCAACC				
2072				
583				
642				
GCAAACTCCCTCGACATGATATCCCTTTTGACCCCAAGCTCTGCTCCCTGACCAACC				
2073				
ATGCCCTTTCCCTTTAACTTCTCAAAACAGATACACAGGGCTTAACTGCTTTACCTCCCTC				
2132				
582				
ATGCCCTTTCCCTTTAACTTCTCAAAACAGATACACAGGGCTTAACTGCTTTACCTCCCTC				
523				
2133				
CTACTGAGTCAGGTTAGGTTGGGAGGTACCCATTTCCGAGTTTAAACCAATGCAATAT				
2192				
522				
CTACTGAGTCAGGTTAGGTTGGGAGGTACCCATTTCCGAGTTTAAACCAATGCAATAT				
463				
2193				
GAGTAAACAAAGTCTATGCTGCTGGGTAGAGAGAGGGGTAGCAAGTTTCATGCTG				
2252				
462				
GAGTAAACAAAGTCTATGCTGCTGGGTAGAGAGAGGGGTAGCAAGTTTCATGCTG				
403				
2253				
TCCTCTTGGTCAATATCTCCAAAGCTCTGATCCCTGCCATGGAAAGTGACACAGGAAA				
2312				
402				
TCCTCTTGGTCAATATCTCCAAAGCTCTGATCCCTGCCATGGAAAGTGACACAGGAAA				
343				
2313				
CATGAGGTCTATGACCTGACAGGATCTTTACTGACGCTCTGCCGCTCGAGGGGAGAGG				
2372				
342				
CATGAGGTCTATGACCTGACAGGATCTTTACTGACGCTCTGCCGCTCGAGGGGAGAGG				
283				
2373				
GGGAGAAGAAGTATGCTGCTGCACATTTCTGAGGCTACTGCAATTTGCTTCAAGGCAGAA				
2432				
282				
GGGAGAAGAAGTATGCTGCTGCACATTTCTGAGGCTACTGCAATTTGCTTCAAGGCAGAA				
223				
2433				
ATCTTCTCTGACGACGCTGACGCTCCAGTTTGGGCGCGATTAAGGAAGTTCTCCGTGCC				
2492				
222				
ATCTTCTCTGACGACGCTGACGCTCCAGTTTGGGCGCGATTAAGGAAGTTCTCCGTGCC				
163				
2493				
TCCTCAGGACGACGAGGAGGAGGCTGACATTTGCGAGTCTCTCTGCGGGCCCAAGGCAG				
2552				
162				
TCCTCAGGACGACGAGGAGGAGGCTGACATTTGCGAGTCTCTCTGCGGGCCCAAGGCAG				
103				
2553				
GTTCAGGAGATCCATCCATAGACGCTCTGGGCGCTCTGCAATTTGCTTCAAGTTTCAGAA				
2612				
102				
GTTCAGGAGATCCATCCATAGACGCTCTGGGCGCTCTGCAATTTGCTTCAAGTTTCAGAA				
43				
2613				
TTAACTGCAAGTATTTTGGAAAGCAAAAAAAGAAAAA				
2654				
42				
TTAACTGCAAGTATTTTGGAAAGCAAAAAAAGAAAAA				
1				
RESULT 9				
BM981939				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				
BASE COUNT				
ORIGIN				
Query Match				
Best Local Similarity				
Matches 702;				
Conservative				
Mismatches				
Indels				
Gaps				
0;				
1953				
ATGTGTTGTGAAAAAAGAGAAATCCCTGGCTCTCGCTGGAGTGGTGGGAGACAAGATTAA				
2012				
643				
702				
ATGTGTTGTGAAAAAAGAGAAATCCCTGGCTCTCGCTGGAGTGGTGGGAGACAAGATTAA				
2013				
GCAAACTCCCTCGACATGATATCCCTTTTGACCCCAAGCTCTGCTCCCTGACCAACC				
2072				
583				
642				
GCAAACTCCCTCGACATGATATCCCTTTTGACCCCAAGCTCTGCTCCCTGACCAACC				
2073				
ATGCCCTTTCCCTTTAACTTCTCAAAACAGATACACAGGGCTTAACTGCTTTACCTCCCTC				
2132				
582				
ATGCCCTTTCCCTTTAACTTCTCAAAACAGATACACAGGGCTTAACTGCTTTACCTCCCTC				
523				
2133				
CTACTGAGTCAGGTTAGGTTGGGAGGTACCCATTTCCGAGTTTAAACCAATGCAATAT				
2192				
522				
CTACTGAGTCAGGTTAGGTTGGGAGGTACCCATTTCCGAGTTTAAACCAATGCAATAT				
463				
2193				
GAGTAAACAAAGTCTATGCTGCTGGGTAGAGAGAGGGGTAGCAAGTTTCATGCTG				
2252				
462				
GAGTAAACAAAGTCTATGCTGCTGGGTAGAGAGAGGGGTAGCAAGTTTCATGCTG				
403				
2253				
TCCTCTTGGTCAATATCTCCAAAGCTCTGATCCCTGCCATGGAAAGTGACACAGGAAA				
2312				
402				
TCCTCTTGGTCAATATCTCCAAAGCTCTGATCCCTGCCATGGAAAGTGACACAGGAAA				
343				
2313				
CATGAGGTCTATGACCTGACAGGATCTTTACTGACGCTCTGCCGCTCGAGGGGAGAGG				
2372				
342				
CATGAGGTCTATGACCTGACAGGATCTTTACTGACGCTCTGCCGCTCGAGGGGAGAGG				
283				
2373				
GGGAGAAGAAGTATGCTGCTGCACATTTCTGAGGCTACTGCAATTTGCTTCAAGGCAGAA				
2432				
282				
GGGAGAAGAAGTATGCTGCTGCACATTTCTGAGGCTACTGCAATTTGCTTCAAGGCAGAA				
223				
2433				
ATCTTCTCTGACGACGCTGACGCTCCAGTTTGGGCGCGATTAAGGAAGTTCTCCGTGCC				
2492				
222				
ATCTTCTCTGACGACGCTGACGCTCCAGTTTGGGCGCGATTAAGGAAGTTCTCCGTGCC				
163				
2493				
TCCTCAGGACGACGAGGAGGAGGCTGACATTTGCGAGTCTCTCTGCGGGCCCAAGGCAG				
2552				
162				
TCCTCAGGACGACGAGGAGGAGGCTGACATTTGCGAGTCTCTCTGCGGGCCCAAGGCAG				
103				
2553				
GTTCAGGAGATCCATCCATAGACGCTCTGGGCGCTCTGCAATTTGCTTCAAGTTTCAGAA				
2612				
102				
GTTCAGGAGATCCATCCATAGACGCTCTGGGCGCTCTGCAATTTGCTTCAAGTTTCAGAA				
43				
2613				
TTAACTGCAAGTATTTTGGAAAGCAAAAAAAGAAAAA				
2654				
42				
TTAACTGCAAGTATTTTGGAAAGCAAAAAAAGAAAAA				
1				
RESULT 9				
BM981939				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				
BASE COUNT				
ORIGIN				
Query Match				
Best Local Similarity				
Matches 702;				
Conservative				
Mismatches				
Indels				
Gaps				
0;				
1953				
ATGTGTTGTGAAAAAAGAGAAATCCCTGGCTCTCGCTGGAGTGGTGGGAGACAAGATTAA				
2012				
643				
702				
ATGTGTTGTGAAAAAAGAGAAATCCCTGGCTCTCGCTGGAGTGGTGGGAGACAAGATTAA				
2013				
GCAAACTCCCTCGACATGATATCCCTTTTGACCCCAAGCTCTGCTCCCTGACCAACC				
2072				
583				
642				
GCAAACTCCCTCGACATGATATCCCTTTTGACCCCAAGCTCTGCTCCCTGACCAACC				
2073				
ATGCCCTTTCCCTTTAACTTCTCAAAACAGATACACAGGGCTTAACTGCTTTACCTCCCTC				
2132				
582				
ATGCCCTTTCCCTTTAACTTCTCAAAACAGATACACAGGGCTTAACTGCTTTACCTCCCTC				
523				
2133				
CTACTGAGTCAGGTTAGGTTGGGAGGTACCCATTTCCGAGTTTAAACCAATGCAATAT				
2192				
522				
CTACTGAGTCAGGTTAGGTTGGGAGGTACCCATTTCCGAGTTTAAACCAATGCAATAT				
463				
2193				
GAGTAAACAAAGTCTATGCTGCTGGGTAGAGAGAGGGGTAGCAAGTTTCATGCTG				
2252				
462				
GAGTAAACAAAGTCTATGCTGCTGGGTAGAGAGAGGGGTAGCAAGTTTCATGCTG				
403				
2253				
TCCTCTTGGTCAATATCTCCAAAGCTCTGATCCCTGCCATGGAAAGTGACACAGGAAA				
2312				
402				
TCCTCTTGGTCAATATCTCCAAAGCTCTGATCCCTGCCATGGAAAGTGACACAGGAAA				
343				
2313				
CATGAGGTCTATGACCTGACAGGATCTTTACTGACGCTCTGCCGCTCGAGGGGAGAGG				
2372				
342				
CATGAGGTCTATGACCTGACAGGATCTTTACTGACGCTCTGCC				

Plate: LLAM10417 row: k column: 08
High quality sequence stop: 782.

FEATURES
source

Location/Qualifiers
1. .1051
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4520191"
/clone_lib="NIH_MGC_91"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
275 a 287 c 314 g 175 t

BASE COUNT 275 a 287 c 314 g 175 t
ORIGIN

Query Match 26.2%; Score 701.2; DB 12; Length 1051;
Best Local Similarity 99.3%; Pred. No. 8.3e-69;
Matches 725; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 40 ACATGAGCGGCGCAGGCGCGTTCGCCACCGGATGGCGAGGCCCGCGAGCGCTCCC 99
DB 1 ACATGAGCGGCGCAGGCGCGTTCGCCACCGGATGGCGAGGCCCGCGAGCGCTCCC 60
QY 100 CGCCCTGCAGTCGAGCTACGACCTCACGGCAAGGTGATGCTTCTGGGAGACACAGGCG 159
DB 61 CGCCCTGCAGTCGAGCTACGACCTCACGGCAAGGTGATGCTTCTGGGAGACACAGGCG 120
QY 160 TCGGCAAAACATGTTTCCTGATCAATTCAAAGACGGGGCTTCCTGTCTGGGAACCTTCA 219
DB 121 TCGGCAAAACATGTTTCCTGATCAATTCAAAGACGGGGCTTCCTGTCTGGGAACCTTCA 180
QY 220 TAGCCACCGTCGCATAGACTTCAGAACAAAGTGTGACTGTGGATGGGTGAGAGTGA 279
DB 181 TAGCCACCGTCGCATAGACTTCAGAACAAAGTGTGACTGTGGATGGGTGAGAGTGA 240
QY 280 AGTGTGAGATCTGGGACACCGCTGGCGAGGAAGGTTCGGAAGCGTCACCATGCTTATT 339
DB 241 AGTGTGAGATCTGGGACACCGCTGGCGAGGAAGGTTCGGAAGCGTCACCATGCTTATT 300
QY 340 ACAGAGATGCTCAGGCTTCCTGCTGCTGATGACATCACCACCAATCTTCTTTGACA 399
DB 301 ACAGAGATGCTCAGGCTTCCTGCTGCTGATGACATCACCACCAATCTTCTTTGACA 360
QY 400 ACATCAGGCGCTGGCTCAGTGCATGATGATGCCAGAGGAGCGTGGTGCATGTC 459
DB 361 ACATCAGGCGCTGGCTCAGTGCATGATGATGCCAGAGGAGCGTGGTGCATGTC 420
QY 460 TGCTAGCAACAAGCGGATATGAGCAGCGAAAGAGTGTCCGTTCCGAGACGGAGAGA 519
DB 421 TGCTAGCAACAAGCGGATATGAGCAGCGAAAGAGTGTCCGTTCCGAGACGGAGAGA 480
QY 520 CCTTGGCCAGGAGTAGGGTTCCTTCTGGAGACCGGCCAGACAGTGGCATGAATG 579
DB 481 CCTTGGCCAGGAGTAGGGTTCCTTCTGGAGACCGGCCAGACAGTGGCATGAATG 540
QY 580 TGGAGTTAGCTTCTTGGCCATGCCAAGAACTGAAATACCGGGCGGGCATCAGCGCG 639
DB 541 TGGAGTTAGCTTCTTGGCCATGCCAAGAACTGAAATACCGGGCGGGCATCAGCGCG 600
QY 640 ATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCAG-AAGAAGCGCTCCAGCTGC 698
DB 601 ATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCAGAAAGAGCGCTCCAGCTGC 660
QY 699 TGCTCTTCATGTGAATCCAGGGGG-CAGAGAGGAGGCTCTGGAGGCACACAGGATGCA 757
DB 661 TGCTCTTCATGTGAATCCAGGGGGCCAGAGAGGAGGCTCTGGAGGCACACAGGATGCA 720
QY 758 GCCTTCCCC 767
DB 721 GCCTTCCCC 730

RESULT 10
BQ028124/c

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE

JOURNAL
COMMENT

FEATURES
source

Location/Qualifiers
1. .700
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3106595"
/clone_lib="NCI_CGAP_Sub9"
/tissue_type="mixed"
/dev_stage="mixed"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73-Pac (Pharmacia) with a modified
polylinker; Site:1: EcoR I; Site:2: Not I; tissues:
Cholonic mucosa with Crohns disease, Cholonic mucosa with
ulcerative colitis, Fetal thymus, Cervix, Cervical
adenosquamous carcinoma, Ligament cells, Prostate
carcinoma, Bladder carcinoma, Brain oligodendroglia ;
NCI_CGAP_Sub9 is a subtracted cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are CGTC, AACG, GGGCC,
GGAAG, TAGC, TAAGC, ATGG, AGACA, ATCAC. For additional
information, contact: Bento Soares, bento-soares@uiowa.edu
TAG_LIB=UI-H-CO0
TAG_TISSUE=Fetal thymus
TAG_SEQ=AACG"

BASE COUNT 166 a 182 c 170 g 182 t
ORIGIN

Query Match 26.1%; Score 696.8; DB 14; Length 700;
Best Local Similarity 99.7%; Pred. No. 3.6e-68;
Matches 698; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1950 GGAATGTGTGAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGACAGAGAT 2009
DB 700 GGAATGTGTGAAAAAAGAGAAATCCCTGGCTCCTGGAGCTGGTGGGACAGAGAT 641
QY 2010 TAAGCAAACTCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCTCCTCCCTGACCA 2069
DB 640 TAAGCAAACTCCCTGACATGTATCCCTTTGACCCCAAGCTCTGCTCCTCCCTGACCA 581

```
QY 2070 CCCATGCCCTTTCCTTAACTTCTCAACAGATACACAGGCGCTAAACTGCTTTACTCCTCC 2129
|||||
Db 580 CCCATGCCCTTTCCTTAACTTCTCAACAGATACACAGGCGCTAAACTGCTTTACTCCTCC 521
QY 2130 CTCCTACTGAGTCAGGTAGGTGGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAA 2189
|||||
Db 520 CTCCTACTGAGTCAGGTAGGTGGGGAGGTCACCCATTTCCGAGTTAAACCAATGCAA 461
QY 2190 TATGAGTAAACAAAGTCATGTGGGTATCTGTGGGTAGAGAGGGGTAGCAAGTTTCA 2249
|||||
Db 460 TATGAGTAAACAAAGTCATGTGGGTATCTGTGGGTAGAGAGGGGTAGCAAGTTTCA 401
QY 2250 GTGTCTCTCTTGGTCACATATCTCCAAAGCTCTGATCCCTGCGCATGGGAAGTGGACAGG 2309
|||||
Db 400 GTGTCTCTCTTGGTCACATATCTCCAAAGCTCTGATCCCTGCGCATGGGAAGTGGACAGG 341
QY 2310 AAACATGAGGTGATGACCTGAGGAGGATCTTTACTGAGCTCTGCGGCGCTGAGGGGGAG 2369
|||||
Db 340 AAACATGAGGTGATGACCTGAGGAGGATCTTTACTGAGCTCTGCGGCGCTGAGGGGGAG 281
QY 2370 AGGGGAGGAAGATGATGCGTGCACATTTCTGAGGCTACTGCAATTTGCTTTCAAGGCA 2429
|||||
Db 280 AGGGGAGGAAGATGATGCGTGCACATTTCTGAGGCTACTGCAATTTGCTTTCAAGGCA 221
QY 2430 GAAATCTTCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCGCGATAGGAAGTTTCCCGTG 2489
|||||
Db 220 GAAATCTTCTCTGAGCAGTCAGCGGCTCCAGTTTGGGCGCGATAGGAAGTTTCCCGTG 161
QY 2490 GCGTCCCTCAGGAGAGGAGGAGGCTGACATTCGAGTCTGTCGCGGCGCGCAAGG 2549
|||||
Db 160 GCGTCCCTCAGGAGAGGAGGAGGCTGACATTCGAGTCTGTCGCGGCGCGCAAGG 101
QY 2550 CAGGTTCAGGAGATCCATCCATAGACAGCTCTGCGGCTCTTSCATTTGAGTTTCA 2609
|||||
Db 100 CAGGTTCAGGAGATCCATCCATAGACAGCTCTGCGGCTCTTSCATTTGAGTTTCA 41
QY 2610 GAATTAACCTGCGATGTTTGGAAAGCAAAAAA 2649
|||||
Db 40 GAATTAACCTGCGATGTTTGGAAAGCAAAAAA 1
```

```
RESULT 11
LOCUS BG282782 1069 bp mRNA linear EST 21-FEB-2001
DEFINITION 502405867F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4518136 5',
mRNA sequence.
ACCESSION BG282782
VERSION BG282782.1 GI:13031999
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1069)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10412 row: e column: 17
High quality sequence stop: 689.
Location/Qualifiers
1. .1069
/organism="Homo sapiens"
/db_xref="taxon:9606"
```

FEATURES source

```
/clone="IMAGE:4518136"
/clone_lib="NIH_MGC_91"
/tissue_type="adenocarcinoma, cell_line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 264 a 296 c 295 g 214 t
ORIGIN
```

```
Query Match 24.1%; Score 643.6; DB 12; Length 1069;
Best Local Similarity 94.3%; Pred. No. 1.7e-62;
Matches 722; Conservative 0; Mismatches 39; Indels 5; Gaps 5;
QY 1104 GGGCAGGTAACTGCTCCAGGAGCTTATCTTCGCCCATCTCCCAATAAGTGGG 1223
|||||
Db 1 GGGCAGGTAACTGCTCCAGGAGCTTATCTTCGCCCATCTCCCAATAAGTGGG 60
QY 1164 GCCTTCACTGTCTCCGCTCTCCAGGAGCTTATCTTCGCCCATCTCCCAATAAGTGGG 1223
|||||
Db 61 GCCTTCACTGTCTCCGCTCTCCAGGAGCTTATCTTCGCCCATCTCCCAATAAGTGGG 120
QY 1224 CCCTTGTGCTGTGAGGAGACCAAGCCCTCAGGAGATATGAGATATGAGATGGAGG 1283
|||||
Db 121 CCCTTGTGCTGTGAGGAGACCAAGCCCTCAGGAGATATGAGATATGAGATGGAGG 180
QY 1284 GGGAGACAAAGGGGAGAGAGTGGTCTAGCTGGCTATCTCTGGCCTTACTTAACACCC 1343
|||||
Db 181 GGGAGACAAAGGGGAGAGAGTGGTCTAGCTGGCTATCTCTGGCCTTACTTAACACCC 240
QY 1344 CTTGAGGCGATGCCCTTTCTCCAGCACACAGACATTTGGGGACCTTGGAAATATGG 1403
|||||
Db 241 CTTGAGGCGATGCCCTTTCTCCAGCACACAGACATTTGGGGACCTTGGAAATATGG 300
QY 1404 TTCCAGGCTCTCTGTTCTCTGGACTTCAGATCTCTGGGGAGCCCCCTCCCTGAATCC 1463
|||||
Db 301 TTCCAGGCTCTCTGTTCTCTGGACTTCAGATCTCTGGGGAGCCCCCTCCCTGAATCC 359
QY 1464 CTGGCTTAGCTACCTTCCTGTGCACCTTAAACACCTCAGTCAAGAAAGAA 1523
|||||
Db 360 CTGGCTTAGCTACCTTCCTGTGCACCTTAAACACCTCAGTCAAGAAAGAA 419
QY 1524 GTTTGTTTATTTTGAATGGAGTCTCGTCTGTCGCCCGAGCTGAGTGCAGTA 1583
|||||
Db 420 GTTATGTTG-TATTTTGAATGGAGTCTCGTCTGTCGCCCGAGCTGAGTGCAGTA 478
QY 1584 GTGCAATCTCCGCTCAGTCAACCTCCCTCCCTGGGGTCAAGCGATCTCCACCTCA 1643
|||||
Db 479 GTGCAATCTCCGCTCAGTCAACCTCCCTCCCTGGGGTCAAGCGATCTCCACCTCA 538
QY 1644 GCGCGGAAGTGTGGGACTATAGTGTGTACCATCACACCTGGCTTAATTTTGTATTT 1703
|||||
Db 539 GCGCGGAAGTGTGGGACTATAGTGTGTACCATCACACCTGGCTTAATTTTGTATTTA 597
QY 1704 TTTGTAGACACAGGTTTCCGCTATGTTGCCAGGCTGTTGAATTTCTGAGCTCAAGC 1763
|||||
Db 598 AAGGTAGACACAGGTTTCCGCTATGTTGCCAGGCTGTTGAATTTCTGAGCTCAAGC 657
QY 1764 AACCTGCCGCGCTCGGCGCTCCCAAGTACTGGGATTTACAGCAG-AAGGCACCATGCCCA 1822
|||||
Db 658 AA-CTGCCGCGCTCGGCGCTCCCAAGTACTGGGATGACCGCGCGAAGGCACATAGCCAG 716
QY 1823 GCGTATAGTGTCTTATCCCAATCTTTTGGCAGGATGAGCTCCA 1868
|||||
Db 717 GCGTATAGTGTCTTATCCCAATCTTTTGGCAGGATGAGCTCCA 762
```

```
RESULT 12
LOCUS BQ183276/c
DEFINITION UI-H-EUO-azr-e-05-0-UI.s1 NCI_CGAP_Carl Homo sapiens cDNA clone
BQ183276
linear EST 30-APR-2002
```

```
IMAGE: 5852524 3', mRNA sequence.
BQ183276
Q0183276.1 GI:20358826
EST.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 629)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapb-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Seq primer: M13 FORWARD
POLYA-Yes
FEATURES
    source
        1..629
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="NCI-CGAP-Carl"
            /tissue_type="Osteoarthritic Cartilage"
            /dev_stage="Adult"
            /lab_host="PH10B (Life Technologies)"
            /notes="Organ: Knee; Vector: pT73-Pac (Pharmacia) with a
            modified polylinker; Site 1: EcoR I; Site 2: Not I;
            NCI-CGAP-Carl is a cDNA library containing the following
            tissue(s): Osteoarthritic Cartilage The library was
            constructed according to Bonaldo, Lennon and Soares,
            Genome Research, 6:791-806, 1996. First strand cDNA
            synthesis was primed with an oligo-dT primer containing a
            Not I site. Double stranded cDNA was ligated to an EcoR I
            adaptor, digested with Not I, and cloned directionally
            into pT73-Pac vector. The oligonucleotide used to prime
            the synthesis of first-strand cDNA contains a library tag
            sequence that is located between the Not I site and the
            (dT)18 tail. The sequence tag for this library is
            TGATCAGCT.
            TAG_L1B=UI-H-EUO
            TAG_TISSUE=osteothritic cartilage
            TAG_SEQ=TCATCACGCT"
BASE COUNT 150 a 162 c 156 g 160 t 1 others
ORIGIN
    Query Match 23.4%; Score 626.4; DB 14; Length 629;
    Best Local Similarity 99.7%; Pred. No. 2.1e-60;
    Matches 627; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2026 GACATGTATCCCTTTGACCCCAAGCTCTGCTCCTGACCAACCATGCTTCCTT 2085
Db 629 GACATGTATCCCTTTGACCCCAAGCTCTGCTCCTGACCAACCATGCTTCCTT 570
Qy 2086 TAATCTTCAACAGATACACAGGCCCTAACTGCTTTACCTCCCTCCTACTGAGTCAGG 2145
Db 569 TAATCTTCAACAGATACACAGGCCCTAACTGCTTTACCTCCCTCCTACTGAGTCAGG 510
Qy 2146 TTAGTGTGGAGGTACCCATTTCCGAGTTAAACCAATATGATGATAAACAAG 2205
Db 509 TTAGTGTGGAGGTACCCATTTCCGAGTTAAACCAATATGATGATAAACAAG 450
Qy 2206 TCATGTGGGTATGTCTGGGGTAGAGAGGGGTAGCAAGTTTCATGTCTCTTGGTCA 2265
Db 449 TCATGTGGGTATGTCTGGGGTAGAGAGGGGTAGCAAGTTTCATGTCTCTTGGTCA 390
Qy 2266 CATATCTCCCAAGCTCTGATCCCTGCCATGGGAGTGGACAGAAACATGAGTCA 2325
Db 389 CATATCTCCCAAGCTCCGATCCCTGCCATGGGAGTGGACAGAAACATGAGTCA 330
Qy 2326 CCTGAGGAGCATCTTTACTGCAGCTCTCGCGCTCGAGGGGAGAGGGAGGAAGT 2385
Db 329 CCTGAGGAGCATCTTTACTGCAGCTCTCGCGCTCGAGGGGAGAGGGAGGAAGT 270
Qy 2386 ATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGCAGCAATCTTGTCTGAG 2445
Db 269 ATGCGCTGCACATTTCTGAGGCTACTGCATTTGCTTTCAAGCAGCAATCTTGTCTGAG 210
Qy 2446 CAGTACAGCGCTCCAGTTTGGGCCGATAAGAGTTCTCCGCTGCCCTCCCTCAGGCAGA 2505
Db 209 CAGTACAGCGCTCCAGTTTGGGCCGATAAGAGTTCTCCGCTGCCCTCCCTCAGGCAGA 150
Qy 2506 CGAGGAGGAGGCTGCATTTGCGAGTCTCTTCTGGGGCCCAAGGAGGAGTTCAGGAGATC 2565
Db 149 CGAGGAGGAGGCTGCATTTGCGAGTCTCTTCTGGGGCCCAAGGAGGAGTTCAGGAGATC 90
Qy 2566 CAATCCCATAGACAGCTCTGGGCTCTTCATTTGAGTTTTCAGAAATTAACCTGCAGTA 2625
Db 89 CAATCCCATAGACAGCTCTGGGCTCTTCGATTTGAGTTTTCAGAAATTAACCTGCAGTA 30
Qy 2626 TTTTGAAGAGCAAAAAAAAAAAAAAAAAAAAA 2654
Db 29 TTTTGAAGAGCAAAAAAAAAAAAAAAAAAAAA 1
RESULT 13
AF188522 607 bp mRNA linear EST 02-MAR-2000
LOCUS AF188522 Homo sapiens ATCC HTB-12; SW1088 Homo sapiens cDNA clone
DEFINITION ISG 3, mRNA sequence.
ACCESSION AF188522
VERSION AF188522.1 GI:7144568
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 607)
Ye, Z. and Connor, J. R.
Identification of Iron Regulated Genes by Rescreening cDNA
Libraries from SSH with Antisense Probe from Three Iron Conditions
Unpublished (2000)
Contact: Ye Z
Neuroscience and Anatomy
Pennsylvania State University College of Medicine
500 University Drive, Hershey, PA 17033, USA
Library screened by SSH and reverse Northern blot; increased
expression in iron loading was confirmed by Northern blot.
FEATURES
    source
        1..607
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="ISG 3"
            /tissue_type="astrocytoma"
            /cell_line="ATCC HTB-12; SW1088"
            /note="Organ: liver"
BASE COUNT 154 a 153 c 155 g 145 t
ORIGIN
    Query Match 22.6%; Score 605.4; DB 9; Length 607;
    Best Local Similarity 99.8%; Pred. No. 4.4e-58;
    Matches 606; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1794 GGGATTACGCGAGAGGACCATGCCAGGCTAGATGTCTTATCCCAATCCTTTGGC 1853
Db 1 GGGATTACGCGAGAGGACCATGCCAGGCTAGATGTCTTATCCCAATCCTTTGGC 60
Qy 1854 AGCATCAGCTCCACAGGCGATTCTTCAAGCAGCTGAAAGTGTTCAGCCCTCCTGGGT 1913
```


DEFINITION 602366910F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4475067 5',
mRNA sequence.
ACCESSION BG253976
VERSION BG253976.1 GI:12763792
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 691)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10300 row: c column: 04
High quality sequence stop: 650.
Location/Qualifiers
i. .691
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4475067"
/clone_lib="NIH_MGC_91"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

FEATURES
source

BASE COUNT 175 a 185 c 190 g 141 t
ORIGIN
Query Match 22.5%; Score 600.6; DB 12; Length 691;
Best Local Similarity 97.6%; Pred. No. 1.3e-57;
Matches 663; Conservative 0; Mismatches 9; Indels 7; Gaps 5;
QY 287 GATCTGGACACCGCTGGCGAGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGA 346
Db 14 GATCTGGACACCGCTGGCGAGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGA 73
QY 347 TGCTCAGGCCCTTCTCTGCTGTATGACATCACCACAAATCTTCTTCGACAAACATCAG 406
Db 74 TGCTCAGGCCCTTCTCTGCTGTATGACATCACCACAAATCTTCTTCGACAAACATCAG 133
QY 407 GGCTGGCTCAGTCAGATTCATGATGCCAGAGGAGCGTGTGATCATGCTGCTAGG 466
Db 134 GGCTGGCTCAGTCAGATTCATGATGCCAGAGGAGCGTGTGATCATGCTGCTAGG 193
QY 467 CAACAAGCGGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGACCTTGGC 526
Db 194 CAACAAGCGGGATATGAGCAGCGAAAGAGTGATCCGTTCCGAAGACGGAGACCTTGGC 253
QY 527 CAGGGAGTAGGGTTCCCTTCTCGAGACCGCCGCAAGACTGGGATGAATGTGGAGTT 586
Db 254 CAGGGAGTAGGGTTCCCTTCTCGAGACCGCCGCAAGACTGGGATGAATGTGGAGTT 313
QY 587 AGCTTTCTGGCCATCGCCAAAGAACTGAATACCGGGCGGCATCAGCGGATGAGCC 646
Db 314 AGCTTTCTGGCCATCGCCAAAGAACTGAATACCGGGCGGCATCAGCGGATGAGCC 373
QY 647 CAGCTTCCAGATCCGAGACTATATAGAGTCCCAAGAGAGCGCTCCAGCTGCTGCTCCTT 706
Db 374 CAGCTTCCAGATCCGAGACTATATAGAGTCCCAAGAGAGCGCTCCAGCTGCTGCTCCTT 433
QY 707 CATGTGATCCCAAGGGGCGAGAGGAGGCTCTGGAGGCACACAGGATGAGCCTTCCCC 766

Db 434 CATGTGAATCCCAAGGGGCGAGAGGAGGCTCTGGAGGCACACAGGATGCAG-CTTCCCC 492
QY 767 CTCCCAGGGCCTGGCTTATTCCAAGAGCTGAGCCAATGGGA--GAAAGATGGAGGACTC 824
Db 493 CTCCCAGGGCCTGGCTTATTCCAAGAGCTGAGCCAATGGGA--GAAAGATGGAGGACTC 552
QY 825 ACTGCACAGCGCTTCTTAGCAGGAGCTATACTCCAACCTCCTTACTTGTAGTTCTCTGC-GG 883
Db 553 ACTGCACAGCGCTTCTTAGCAGGAGCTATACTCCAACCTCCTTACTTGTAGTTCTCTGCAGG 612
QY 884 TCTCCCGCGCATCCACAGGGA-GGGTAAACACTTAGCTTTTATTTT--AATAGTACATAA 940
Db 613 TCTCCCGCGCATCCACAGGAGGGGTAAACACTTAGCTTTTATTTTCTCAATACGTACATAA 672
QY 941 TTTAATACCAAAAAAGGCG 959
Db 673 TTTAATACCAAAAAAGGCG 691

Search completed: November 17, 2002, 16:36:29
Job time : 2452 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 17, 2002, 13:13:23 ; Search time 4598 Seconds
(without alignments)
16924.946 Million cell updates/sec

Title: US-09-817-199a-1
Perfect score: 2674
Sequence: 1 ttccctgcgcggcgccact.....aaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_on.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2161.8	80.8	2195	9	AK057069	AK057069 Homo sapi
2	1562	73.4	1977	9	AK098068	AK098068 Homo sapi
3	1354.8	50.7	3105	9	AK054846	AK054846 Homo sapi
4	1086	40.6	1116	6	AX236082	AX236082 Sequence
5	938.4	35.1	1700	9	BC016615	BC016615 Homo sapi
6	576	21.5	576	6	AX236084	AX236084 Sequence
7	563.6	21.1	690	10	AF233582	AF233582 Mus muscu
8	315.6	11.8	1513	9	BC007681	BC007681 Homo sapi
9	312.4	11.7	1098	10	RNU18771	U18771 Rattus norv
10	304.6	11.4	1320	9	AB027137	AB027137 Homo sapi
11	303.2	11.3	573	9	AF498952	AF498952 Homo sapi
12	292.6	10.9	1340	6	AR062279	AR062279 Sequence
13	234.2	8.8	2536	3	AY061826	AY061826 Drosophil
14	214.8	8.0	127587	9	HS941P9	Z95331 Human DNA s
15	214.2	8.0	102332	9	AC010352	AC010352 Homo sapi
16	214.2	8.0	107465	9	AC010434	AC010434 Homo sapi
17	212.8	8.0	146740	9	CNS01D7X	AL132989 Human chr
18	212.6	8.0	174309	9	AC016656	AC016656 Homo sapi
19	212.6	8.0	177640	9	AC016652	AC016652 Homo sapi
20	211.6	7.9	148841	9	AC011462	AC011462 Homo sapi
21	210.4	7.9	199601	9	AL389915	AL389915 Human DNA
22	210.2	7.9	90000	9	AC117467	AC117467 Homo sapi
23	209.6	7.8	141771	9	AC091821	AC091821 Homo sapi
24	209	7.8	110000	2	AC013391_0	AC013391 Homo sapi
25	209	7.8	125980	9	AC093223	AC093223 Homo sapi
26	209	7.8	157141	9	AC011470	AC011470 Homo sapi
27	209	7.8	176304	9	AC013787	AC013787 Homo sapi
28	208.4	7.8	167509	2	AC012107	AC012107 Homo sapi
29	208	7.8	43790	9	AC113388	AC113388 Homo sapi
30	208	7.8	86685	9	AC118457	AC118457 Homo sapi
31	208	7.8	107995	2	HSJ356711	AL121971 Homo sapi
32	208	7.8	172048	9	HS179N16	Z95152 Homo sapien
33	207.6	7.8	132721	9	AC004552	AC004552 Homo sapi
34	207.6	7.8	207366	9	AC090001	AC090001 Homo sapi
35	207.4	7.8	136222	9	AC004099	AC004099 Homo sapi
36	207.2	7.7	63905	9	AC022515	AC022515 Homo sapi
37	207.2	7.7	186330	9	AC007376	AC007376 Homo sapi
38	207.2	7.7	271814	2	AC027406	AC027406 Homo sapi
39	207	7.7	154289	2	AC027714	AC027714 Homo sapi
40	207	7.7	161757	2	AC044901	AC044901 Homo sapi
41	207	7.7	163528	2	AC130416	AC130416 Homo sapi
42	207	7.7	187399	9	AC090970	AC090970 Homo sapi
43	207	7.7	209797	2	AC005748	AC005748 Homo sapi
44	207	7.7	346940	2	AC107420	AC107420 Homo sapi
45	206.8	7.7	157113	9	AC083797	AC083797 Homo sapi

ALIGNMENTS

RESULT 1
AK057069
LOCUS
DEFINITION Homo sapiens cDNA FLJ32507 fis, clone SMINT1000048, moderately similar to Mus musculus GTPase Rab37 mRNA.
ACCESSION AK057069
VERSION AK057069.1 GI:16552648
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens small intestine cDNA to mRNA, clone_lib:SMINT1 clone:SMINT1000048.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,

Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K.
and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2195)
Isogai,T., Otsuki,T. and Sugiyama,T.
Direct Submission
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: RAB and
HRI.

TITLE NEDO human cDNA sequencing project

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2195)

AUTHORS Isogai,T., Otsuki,T. and Sugiyama,T.

TITLE Direct Submission

JOURNAL

COMMENT

FEATURES

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="SMINT1000048"

/tissue.type="small intestine"

/clone.lib="SMINT1"

/note="cloning vector: pME185FL3"

511 a 596 c 582 g 506 t

BASE COUNT

ORIGIN

Query Match 80.8%; Score 2161.8; DB 9; Length 2195;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2163; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 472 AGCGGATATGACGAGCGAAGAGTAGTCGGTTCGGAAGACGGAGACGCTTGCCAGGG 531

DB 31 AGCGGATATGACGAGCGAAGAGTAGTCGGTTCGGAAGACGGAGACGCTTGCCAGGG 90

QY 532 AGTACGGTGTTCCTTCCCTGGAGACCGCCGAGACTGGCATGAATGTGGAGTTAGCCT 591

DB 91 AGTACGGTGTTCCTTCCCTGGAGACCGCCGAGACTGGCATGAATGTGGAGTTAGCCT 150

QY 592 TTCTGGCCATCGCCAGGAAGTGAATATACCGGGCGGGCATGAGGGGATGAGCCAGCT 651

DB 151 TTCTGGCCATCGCCAGGAAGTGAATATACCGGGCGGGCATGAGGGGATGAGCCAGCT 210

QY 652 TCCAGATCCGAGACTATGTAGAGTCCAGAGAGCGCTCCAGCTGCTGCTCTTCATGT 711

DB 211 TCCAGATCCGAGACTATGTAGAGTCCAGAGAGCGCTCCAGCTGCTGCTCTTCATGT 270

QY 712 GAATCCAGGGGGCAGAGAGGAGCTCTGGAGCACACAGGATGCCCTTCCCTCC 771

DB 271 GAATCCAGGGGGCAGAGAGGAGCTCTGGAGCACACAGGATGCCCTTCCCTCC 330

QY 772 AGGCCTGGCTTATTCGAAGAGCTTGACCAATGGGGAAGAAGATGGAGGACTCACTGCAC 831

DB 331 AGGCCTGGCTTATTCGAAGAGCTTGACCAATGGGGAAGAAGATGGAGGACTCACTGCAC 390

QY 832 AGCCGCTTCTAGCAGGAGGACTATCTCCAACTCTCTACTTGGTTCCTGGGCTCCCG 891

DB 391 AGCCGCTTCTAGCAGGAGGACTATCTCCAACTCTCTACTTGGTTCCTGGGCTCCCG 450

QY 892 CATCCAGGGAGGTAACACATTAGCTTTATTTTAAATAGTACATAATTTAATACAA 951

DB 451 CATCCAGGGAGGTAACACATTAGCTTTATTTTAAATAGTACATAATTTAATACAA 510

QY 952 AAAAGGCGCTTGATCCCAAAAAACCGAGCTGGAGCTAGTGGCCCTTTGCTTTCTA 1011

DB 511 AAAAGGCGCTTGATCCCAAAAAACCGAGCTGGAGCTAGTGGCCCTTTGCTTTCTA 570

QY 1012 GGACTTTGGGGGGCGGCGCTCCCTTAAGCATACAAAGCTGGTGTGCTCCAGCTCAG 1071

DB 571 GGACTTTGGGGGGCGGCGCTCCCTTAAGCATACAAAGCTGGTGTGCTCCAGCTCAG 630

QY 1072 CCCAGGGGACACACATGACATCTTTGGGGGTGAGGCGAGTAACTACTCCATCGCACCTC 1131

DB 631 CCCAGGGGACACACATGACATCTTTGGGGGTGAGGCGAGTAACTACTCCATCGCACCTC 690

QY 1132 AGTTCCAGCTGGACAGAGGCTCAGGTGACCCCGACCTTCACTGTCTCCCGCTCTCCAGGAG 1191

DB 691 AGTTCCAGCTGGACAGAGGCTCAGGTGACCCCGACCTTCACTGTCTCCCGCTCTCCAGGAG 750

QY 1192 CTTATCTTCGCCCCATCTCCCAATAAGTGGCCCTTGTCTGTGAGGAAGACCAAGCC 1251

DB 751 CTTATCTTCGCCCCATCTCCCAATAAGTGGCCCTTGTCTGTGAGGAAGACCAAGCC 810

QY 1252 TCAGGAAAGATAAGAGATATGAGATGGAGGGGAGGACAAAGGGCAGAGTAGGGTC 1311

DB 811 TCAGGAAAGATAAGAGATATGAGATGGAGGGGAGGACAAAGGGCAGAGTAGGGTC 870

QY 1312 TAGCTGGCTATCTCTGGCCTTACTAACACCCCTTGAGGCGATGCCCTTTTCTCCAGCA 1371

DB 871 TAGCTGGCTATCTCTGGCCTTACTAACACCCCTTGAGGCGATGCCCTTTTCTCCAGCA 930

QY 1372 CACAAGCACATTTGGGCACCTGGAAATATTTGGTTCAGGCTCCTGTCTCTGGACTTCAG 1431

DB 931 CACAAGCACATTTGGGCACCTGGAAATATTTGGTTCAGGCTCCTGTCTCTGGACTTCAG 990

QY 1432 ATCTTGGGGAGCCCTCCCGCCCTGAATCCCTGGCTTAGCTTCTCCCTCTGCA 1491

DB 991 ATCTTGGGGAGCCCTCCCGCCCTGAATCCCTGGCTTAGCTTCTCCCTCTGCA 1050

QY 1492 CTTAAAACTCAGGTCAGAACTAGAAAAAGATTTGTTTTTATTTTGGAAATGGAG 1551

DB 1051 CTTAAAACTCAGGTCAGAACTAGAAAAAGATTTGTTTTTATTTTGGAAATGGAG 1110

QY 1552 TCTCGTCTGTGCCAGGCTGAGTGCAGTAGTGCATCTCGCTCACTACAACCTCCA 1611

DB 1111 TCTCGTCTGTGCCAGGCTGAGTGCAGTAGTGCATCTCGCTCACTACAACCTCCA 1170

QY 1612 CTCCCTGGGGCTCAAGCGATCTCCACCTCAGCCCGGAGTAGCTGGAGCTATAGGTG 1671

DB 1171 CTCCCTGGGGCTCAAGCGATCTCCACCTCAGCCCGGAGTAGCTGGAGCTATAGGTG 1230

QY 1672 TGTACCATCACCTGGCTAAATTTTGTATTTTGTACACAGGTTTCGCCATGTTG 1731

DB 1231 TGTACCATCACCTGGCTAAATTTTGTATTTTGTACACAGGTTTCGCCATGTTG 1290

QY 1732 CCCAGCTGTCTTGAATTCCTGAGCTCAAGCAACCTCGCGCTCGGCTCCCAAGTA 1791

DB 1291 CCCAGCTGTCTTGAATTCCTGAGCTCAAGCAACCTCGCGCTCGGCTCCCAAGTA 1350

QY 1792 CTGGGATTACACGAGAAGGCACATGCCAGCTAGATGTGTTCATCCCAATCCTTTG 1851

DB 1351 CTGGGATTACACGAGAAGGCACATGCCAGCTAGATGTGTTCATCCCAATCCTTTG 1410

QY 1852 GCAGGATCAGCTCCACAGGCGATTTCTTCAAGAGCTGAAGTGTTCAGCTCCTGGG 1911

DB 1411 GCAGGATCAGCTCCACAGGCGATTTCTTCAAGAGCTGAAGTGTTCAGCTCCTGGG 1470

QY 1912 TTAAGAGCCAGATAAGGAGAAATCCCTTTCTAGTGTTCGAATGTGTGTGAAAAAAG 1971

DB 1471 TTAAGAGCCAGATAAGGAGAAATCCCTTTCTAGTGTTCGAATGTGTGTGAAAAAAG 1530

QY 1972 AGAAATCCCTGGCTCCTGGAGCTGGGGAGACAAGATAAGCAAACTCCCTCGACATG 2031

DB 1531 AGAAATCCCTGGCTCCTGGAGCTGGGGAGACAAGATAAGCAAACTCCCTCGACATG 1590

QY 2032 TATCCCTTTGACCCCAAGCTCTGCCCTCCCTGACCCCACTGCTTCTTAACTT 2091

DB 1591 TATCCCTTTGACCCCAAGCTCTGCCCTCCCTGACCCCACTGCTTCTTAACTT 1650

QY 2092 CTCAAGAGATACCGGGCTTAACTGCTTTACCTCCCTCTCTACTGAGTCAGGTAGGT 2151

```
|||||
Db 1651 CTCAACAGATACACAGGCGCTAACTGCTTTACCTCCCTCTACTGAGTCAGGTAGGT 1710
QY 2152 GGTGGAGGTACCATTCCTCGAGTTAAACCAATGCAATATGAGTAAACAAGTCATGT 2211
Db 1711 GGTGGAGGTACCATTCCTCGAGTTAAACCAATGCAATATGAGTAAACAAGTCATGT 1770
QY 2212 GGTATGTCTGGGTAGAGAGAGGGGTAGCAAGTTTCATGTCTCTCTTGGTTCACATATC 2271
Db 1771 GGTATGTCTGGGTAGAGAGAGGGGTAGCAAGTTTCATGTCTCTCTTGGTTCACATATC 1830
QY 2272 TCCCAAGCTCTGATCTCCCTGCGCATGGGAAGTGGACAGGAACATGAGTTCATGACCTGCA 2331
Db 1831 TCCCAAGCTCCGATCCCTGCGCATGGGAAGTGGACAGGAACATGAGTTCATGACCTGCA 1890
QY 2332 GGCATCTTTACTGACAGCTCTCGCGCTCGAGGGGAGGGGAGGAGGAAGATGATGCGC 2391
Db 1891 GGCATCTTTACTGACAGCTCTCGCGCTCGAGGGGAGGGGAGGAGGAAGATGATGCGC 1950
QY 2392 TGCACATTTCTGAGGCTACTGCTATTTGCTTTCAAGGCGAGAAATCTTGCTCTGAGCAGTCA 2451
Db 1951 TGCACATTTCTGAGGCTACTGCTATTTGCTTTCAAGGCGAGAAATCTTGCTCTGAGCAGTCA 2010
QY 2452 GCGGCTCCAGTTTGGGCCCGGATPAAGGAAGTTCTCCGTGGCCCTCCCTCAGGCGAGACAGGG 2511
Db 2011 GCGGCTCCAGTTTGGGCCCGGATPAAGGAAGTTCTCCGTGGCCCTCCCTCAGGCGAGACAGGG 2070
QY 2512 AGGAGGCTGACATTCAGGCTCTCTTCTGGGGCCCAAGGAGGTTGCGAGGATCAATCC 2571
Db 2071 AGGAGGCTGACATTCAGGCTCTCTTCTGGGGCCCAAGGAGGTTGCGAGGATCAATCC 2130
QY 2572 CATGACAGCTCTGGGCTCTTGCAATTTGAGTTTTCAGAAATTAAGTGCAGTATTTGG 2631
Db 2131 CATGACAGCTCTGGGCTCTTGCAATTTGAGTTTTCAGAAATTAAGTGCAGTATTTGG 2190
QY 2632 AAAGC 2636
|||||
Db 2191 AAAGC 2195

RESULT 2
AK098068
LOCUS
DEFINITION Homo sapiens cDNA FLJ40749 fis, clone TRACH2000540, highly similar
to Mus musculus GTPase Rab37 (Rab37) mRNA.
ACCESSION AK098068
VERSION AK098068.1 GI:21758001
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens trachea cDNA to mRNA, clone_lib=TRACH2
clone:TRACH2000540.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,
Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakamatsu, A.,
Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K.,
Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y.,
Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K.,
Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,
Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
NEO human cDNA sequencing project
Unpublished
2 (bases 1 to 1977)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
```

```
Research Association for Biotechnology (RAB): cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.): 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
Location/Qualifiers
1..1977
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TRACH2000540"
/tissue_type="trachea"
/clone_lib="TRACH2"
/note="cloning vector: pME18SFL3"
57..728
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC05227.1"
/db_xref="GI:21758002"
/translation="MTGTPGAVATRDGEAPERSPPSPSYDLTGKVMLLGDTGVGKTC
FLIOFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGDERFSVTHAYRD
AQALLLLDYDTNKSFDNIRAWLFEIHEYAQRDVIYVLLGNKADMSRSRVSREDEGT
LAREYGVFFLETSKRTGMNVELAFIALAKELKIRAGHQADEPFSQIIRDYVESQKRSS
CCSPM"
BASE COUNT 442 a 560 c 539 g 436 t
ORIGIN
Query Match 73.4%; Score 1962; DB 9; Length 1977;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1962; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTCGGCTCGCGGCGGCGACTGCTCACCTCTCGTCAGGGACATGACGGGCACGCCAGGGG 60
Db 16 TTCGGCTCGCGGCGGCGACTGCTCACCTCTCGTCAGGGACATGACGGGCACGCCAGGGG 75
QY 61 CCGTTGCCAACCCTGGGATGGCGAGGCCCCCGAGCGCTCCCGCCCTGCAGTCGGAGTAGC 120
Db 76 CCGTTGCCAACCCTGGGATGGCGAGGCCCCCGAGCGCTCCCGCCCTGCAGTCGGAGTAGC 135
QY 121 ACCTCAGCGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCTCTCA 180
Db 136 ACCTCAGCGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCTCTCA 195
QY 181 TCCAATTTCAAGACAGGGGCGCTTCTCTCGGGAACCTTCATAGCCACCGTCGGCATAGACT 240
Db 196 TCCAATTTCAAGACAGGGGCGCTTCTCTCGGGAACCTTCATAGCCACCGTCGGCATAGACT 255
QY 241 TCAGGAACAAGGTGCTGACTGTGGATGGCGTGAGAGTGAAGTGCAGATCTGGGACACCG 300
Db 256 TCAGGAACAAGGTGCTGACTGTGGATGGCGTGAGAGTGAAGTGCAGATCTGGGACACCG 315
QY 301 CTGGGCAAGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCGTTGC 360
Db 316 CTGGGCAAGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCGTTGC 375
QY 361 TTCTGCTGTATGACATCACCAACAATCTCTTCGACAACATCAGGCGCTTGCTCAGCTG 420
Db 376 TTCTGCTGTATGACATCACCAACAATCTCTTCGACAACATCAGGCGCTTGCTCAGCTG 435
QY 421 AGATTTCATGAGTATGCCAGAGGAGCGTGGTGATCATGCTGTAGGCAACAAGGCGGATA 480
Db 436 AGATTTCATGAGTATGCCAGAGGAGCGTGGTGATCATGCTGTAGGCAACAAGGCGGATA 495
QY 481 TGAGCAGCAAGAGATGATCCGTTCCGAGAGCGGAGACCTTGGCCAGGGAGTAGCGGTG 540
Db 496 TGAGCAGCAAGAGATGATCCGTTCCGAGAGCGGAGACCTTGGCCAGGGAGTAGCGGTG 555
QY 541 TTCCCTTCTCTGAGACACGCGCAAGACTGGCATGAATGTGGAGTTAGCTTCTTGCGCA 600
Db 556 TTCCCTTCTCTGAGACACGCGCAAGACTGGCATGAATGTGGAGTTAGCTTCTTGCGCA 615
QY 601 TCGCCAAGGAACCTGAAATACCGGGCGGCGCATCAGGCGGATGAGCCAGCTTCCAGATCC 660
|||||
```

Research Association for Biotechnology (RAB): cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.): 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.

FEATURES
Location/Qualifiers
1..1977

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TRACH2000540"
/tissue_type="trachea"
/clone_lib="TRACH2"
/note="cloning vector: pME18SFL3"

57..728
/note="unnamed protein product"

/codon_start=1
/protein_id="BAC05227.1"
/db_xref="GI:21758002"
/translation="MTGTPGAVATRDGEAPERSPPSPSYDLTGKVMLLGDTGVGKTC
FLIOFKDGAFLSGTFIATVGIDFRNKVVTVDGVRVKLQIWDTAGDERFSVTHAYRD
AQALLLLDYDTNKSFDNIRAWLFEIHEYAQRDVIYVLLGNKADMSRSRVSREDEGT
LAREYGVFFLETSKRTGMNVELAFIALAKELKIRAGHQADEPFSQIIRDYVESQKRSS
CCSPM"

BASE COUNT 442 a 560 c 539 g 436 t
ORIGIN

Query Match 73.4%; Score 1962; DB 9; Length 1977;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1962; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTCGGCTCGCGGCGGCGACTGCTCACCTCTCGTCAGGGACATGACGGGCACGCCAGGGG 60
Db 16 TTCGGCTCGCGGCGGCGACTGCTCACCTCTCGTCAGGGACATGACGGGCACGCCAGGGG 75

QY 61 CCGTTGCCAACCCTGGGATGGCGAGGCCCCCGAGCGCTCCCGCCCTGCAGTCGGAGTAGC 120
Db 76 CCGTTGCCAACCCTGGGATGGCGAGGCCCCCGAGCGCTCCCGCCCTGCAGTCGGAGTAGC 135

QY 121 ACCTCAGCGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCTCTCA 180
Db 136 ACCTCAGCGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTTCTCTCA 195

QY 181 TCCAATTTCAAGACAGGGGCGCTTCTCTCGGGAACCTTCATAGCCACCGTCGGCATAGACT 240
Db 196 TCCAATTTCAAGACAGGGGCGCTTCTCTCGGGAACCTTCATAGCCACCGTCGGCATAGACT 255

QY 241 TCAGGAACAAGGTGCTGACTGTGGATGGCGTGAGAGTGAAGTGCAGATCTGGGACACCG 300
Db 256 TCAGGAACAAGGTGCTGACTGTGGATGGCGTGAGAGTGAAGTGCAGATCTGGGACACCG 315

QY 301 CTGGGCAAGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCGTTGC 360
Db 316 CTGGGCAAGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCGTTGC 375

QY 361 TTCTGCTGTATGACATCACCAACAATCTCTTCGACAACATCAGGCGCTTGCTCAGCTG 420
Db 376 TTCTGCTGTATGACATCACCAACAATCTCTTCGACAACATCAGGCGCTTGCTCAGCTG 435

QY 421 AGATTTCATGAGTATGCCAGAGGAGCGTGGTGATCATGCTGTAGGCAACAAGGCGGATA 480
Db 436 AGATTTCATGAGTATGCCAGAGGAGCGTGGTGATCATGCTGTAGGCAACAAGGCGGATA 495

QY 481 TGAGCAGCAAGAGATGATCCGTTCCGAGAGCGGAGACCTTGGCCAGGGAGTAGCGGTG 540
Db 496 TGAGCAGCAAGAGATGATCCGTTCCGAGAGCGGAGACCTTGGCCAGGGAGTAGCGGTG 555

QY 541 TTCCCTTCTCTGAGACACGCGCAAGACTGGCATGAATGTGGAGTTAGCTTCTTGCGCA 600
Db 556 TTCCCTTCTCTGAGACACGCGCAAGACTGGCATGAATGTGGAGTTAGCTTCTTGCGCA 615

QY 601 TCGCCAAGGAACCTGAAATACCGGGCGGCGCATCAGGCGGATGAGCCAGCTTCCAGATCC 660
|||||

Db 616 TCGCCAAGGAACGTGAATAACCGCGCGGCATCAGCGGATGAGCCAGCTTCCAGATCC 675
QY 661 GAGACTATGTAGAGTCCCAAGAAAGCGCTCCAGCTGCTGCTCTTATGTGAATCCCAAG 720
Db 676 GAGACTATGTAGAGTCCCAAGAAAGCGCTCCAGCTGCTGCTCTTATGTGAATCCCAAG 735
QY 721 GGGCGCAGAGAGGAGGCTCTGGAGGCACACAGGATGAGGCTTCCCTCCAGGCGCTGGC 780
Db 736 GGGCGCAGAGAGGAGGCTCTGGAGGCACACAGGATGAGGCTTCCCTCCAGGCGCTGGC 795
QY 781 TTATTCAGAGAGGCTGAGCCATGGGAGAAAGATGAGGACTCACTGCACAGCCCTTC 840
Db 796 TTATTCAGAGAGGCTGAGCCATGGGAGAAAGATGAGGACTCACTGCACAGCCCTTC 855
QY 841 CTAGCAGGAGCTATCTCCAACTCTACTTGTAGTTCCTGCGGTCTCCCGCATCCACAG 900
Db 856 CTAGCAGGAGCTATCTCCAACTCTACTTGTAGTTCCTGCGGTCTCCCGCATCCACAG 915
QY 901 GGAGGGTAAACACTTAGCTTTTATTTTATAGTACATAATTTAATACCAAAAAGGCGC 960
Db 916 GGAGGGTAAACACTTAGCTTTTATTTTATAGTACATAATTTAATACCAAAAAGGCGC 975
QY 961 CTGGATCCCCAAAACCGAGGCTGGAGCTAGTGGCCCTTTTGCCTTCTAGGACTTGG 1020
Db 976 CTGGATCCCCAAAACCGAGGCTGGAGCTAGTGGCCCTTTTGCCTTCTAGGACTTGG 1035
QY 1021 GGGCGGCGCTCCCTCTTAAGCATACAAAGTGGTGTCTCCAGCTCAGCCCGCAGGGG 1080
Db 1036 GGGCGGCGCTCCCTCTTAAGCATACAAAGTGGTGTCTCCAGCTCAGCCCGCAGGGG 1095
QY 1081 ACACAGATGACATTTGGGGTGAGGCGAGTAACTGCTCCATCGACCTTCAGTTCAAGCT 1140
Db 1096 ACACAGATGACATTTGGGGTGAGGCGAGTAACTGCTCCATCGACCTTCAGTTCAAGCT 1155
QY 1141 GGACAGAGGCTCAGTACCCCGCTTCTACTGCTCCCGCTCTCCAGGAGCTTACTTTC 1200
Db 1156 GGACAGAGGCTCAGTACCCCGCTTCTACTGCTCCCGCTCTCCAGGAGCTTACTTTC 1215
QY 1201 GCCCATCTCCCAATAAGTGGGCGCTTGTGCTGTAGGAAAGCAAGGCTCAGGGAAG 1260
Db 1216 GCCCATCTCCCAATAAGTGGGCGCTTGTGCTGTAGGAAAGCAAGGCTCAGGGAAG 1275
QY 1261 ATAAGAGATGAGATGGAGGGGAGGACAAAGGCGCAGAGTAGGGTCTAGCTGGCT 1320
Db 1276 ATAAGAGATGAGATGGAGGGGAGGACAAAGGCGCAGAGTAGGGTCTAGCTGGCT 1335
QY 1321 ATCTCTGCGCTTACTAACACCCCTCGAGGCGATGCCCTTCTCAGCACACAAGCAC 1380
Db 1336 ATCTCTGCGCTTACTAACACCCCTCGAGGCGATGCCCTTCTCAGCACACAAGCAC 1395
QY 1381 ATGGGCGACCTGGAATATTGGTTCCAGGCTCCTGTTCTCTGAGCTTCAGATCCTGGGG 1440
Db 1396 ATGGGCGACCTGGAATATTGGTTCCAGGCTCCTGTTCTCTGAGCTTCAGATCCTGGGG 1455
QY 1441 GAGCCCTCCCGCCCTGAATCCCTGCTTAGCTACCTTCTGCTGCTGACCTTAAAC 1500
Db 1456 GAGCCCTCCCGCCCTGAATCCCTGCTTAGCTACCTTCTGCTGCTGACCTTAAAC 1515
QY 1501 CTCAGGTCAGAACTAGGAAAGAGTTTGTATTTTATTTTGAATGGAGTCTCGTTCT 1560
Db 1516 CTCAGGTCAGAACTAGGAAAGAGTTTGTATTTTATTTTGAATGGAGTCTCGTTCT 1575
QY 1561 GTCGCCAGGCTGAGGTGCAGTAGTCAATCTCCGCTCACTACAACTCCACCTCCCTGGG 1620
Db 1576 GTCGCCAGGCTGAGGTGCAGTAGTCAATCTCCGCTCACTACAACTCCACCTCCCTGGG 1635
QY 1621 GCTCAAGGATCTCCCACTCAGCCCGCAAGTAGTGGACTATAGGTGTGTACCATC 1680
Db 1636 GCTCAAGGATCTCCCACTCAGCCCGCAAGTAGTGGACTATAGGTGTGTACCATC 1695
QY 1681 ACACCTGGCTAAATTTTGTATTTTGTAGACACAGGTTTTCGCCATGTTGCCAGGCTG 1740
Db 1696 ACACCTGGCTAAATTTTGTATTTTGTAGACACAGGTTTTCGCCATGTTGCCAGGCTG 1755

QY 1741 GTCTTGAATTCCTGAGCTCAAGCAACCTGCGGCGCTCGCCCTCCCAAGTACTGGGATTA 1800
Db 1756 GTCTTGAATTCCTGAGCTCAAGCAACCTGCGGCGCTCGCCCTCCCAAGTACTGGGATTA 1815
QY 1801 CAGCAGAAAGGACCATGCCAGGCTAGATGTCTTATTCCTCAATCTTTGGCAGGCATG 1860
Db 1816 CAGCAGAAAGGACCATGCCAGGCTAGATGTCTTATTCCTCAATCTTTGGCAGGCATG 1875
QY 1861 CAGCTCCACAGCGATTCTTCAAGCAGCTGAAGTGTCTTAGCCCTCTGGTTAAAGCC 1920
Db 1876 CAGCTCCACAGCGATTCTTCAAGCAGCTGAAGTGTCTTAGCCCTCTGGTTAAAGCC 1935
QY 1921 AGATAGGAGAAATCCCTTCTAGTGTGGTGGATGTGTGTG 1962
Db 1936 AGATAGGAGAAATCCCTTCTAGTGTGGTGGATGTGTGTG 1977

RESULT 3
AK054846
LOCUS
DEFINITION
Homo sapiens cDNA FLJ30284 fis, clone BRACE2002812, moderately similar to Mus musculus GTPase Rab37 (Rab37) mRNA.
ACCESSION
AK054846
VERSION
AK054846.1 GI:16549463
KEYWORDS
Homo sapiens cerebellum cDNA to mRNA, clone_lib:BRACE2
SOURCE
clone:BRACE2002812.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1
AUTHORS
Nishi,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yanashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fuji,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T., NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 3105)
Isogai,T., Otsuki,T. and Sugiyama,T.
Direct Submission
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES
source
Location/Qualifiers
i. 3105
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="BRACE2002812"
/tissue.type="cerebellum"
/clone_lib="BRACE2"
/note="cloning vector: pME18SFL3"
BASE COUNT 680 a 905 c 898 g 622 t
ORIGIN

Query Match 50.7%; Score 1354.8; DB 9; Length 3105;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1362; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 589 CTTTCTGGCGCTCAGCAGGAACCTGAAATACCGGCGCGGCATCAGCGGATGAGCCCA 648
Db 1732 CCATTGTCTTCTTCAAGGGAACCTGAAATACCGGCGCGGCATCAGCGGATGAGCCCA 1791

QY 649 GCTTCCAGATCCGAGACATATGTAGAGTCCAGAAAGCGCTCCAGCTGCTGCTCTTCA 708
|||||
Db 1792 GCTTCCAGATCCGAGACATATGTAGAGTCCAGAAAGCGCTCCAGCTGCTGCTCTTCA 1851
QY 709 TGTGAATCCAGGGGCGAGAGAGGAGTCTGGAGGCACACAGGATGACAGCTTCCCGCT 768
|||||
Db 1852 TGTGAATCCAGGGGCGAGAGAGGAGTCTGGAGGCACACAGGATGACAGCTTCCCGCT 1911
QY 769 CCCAGGCTGGCTTATCCAGAGGCTGAGCAATGGGGAGAAAGATGAGAGCTACAGT 828
|||||
Db 1912 CCCAGGCTGGCTTATCCAGAGGCTGAGCAATGGGGAGAAAGATGAGAGCTACAGT 1971
QY 829 CACAGCGCTTCTAGCAGGAGCTATATCTCCAACCTCTTACTTGAGTTCTCTGGGCTCTCC 888
|||||
Db 1972 CACAGCGCTTCTAGCAGGAGCTATATCTCCAACCTCTTACTTGAGTTCTCTGGGCTCTCC 2031
QY 889 CCGCATCCACAGGGAGGTTAAACACTTAGCTTTTATTTTAATAGTACATAATTTAATAC 948
|||||
Db 2032 CCGCATCCACAGGGAGGTTAAACACTTAGCTTTTATTTTAATAGTACATAATTTAATAC 2091
QY 949 CAAAAAGCGCTGGATCCCCAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTT 1008
|||||
Db 2092 CAAAAAGCGCTGGATCCCCAAAAACCGAGGCTGGGAGCTAGTGGCCCTTTTGCTTT 2151
QY 1009 CTAGGACTTGGGGGCGGCCCTCCCTCCTTAAGCATATAACAAAGTGGTGTGCTCCAGCT 1068
|||||
Db 2152 CTAGGACTTGGGGGCGGCCCTCCCTCCTTAAGCATATAACAAAGTGGTGTGCTCCAGCT 2211
QY 1069 CAGCCCCAGGGGACACAGATGACATCTTTGGGGTGGGAGGTAATGACTCCATCGCAC 1128
|||||
Db 2212 CAGCCCCAGGGGACACAGATGACATCTTTGGGGTGGGAGGTAATGACTCCATCGCAC 2271
QY 1129 CTACGTTAGCTGGACAGAGGCTCAGGTGACCCAGCCCTCAGTGTCTCCCGCTCCAG 1188
|||||
Db 2272 CTACGTTAGCTGGACAGAGGCTCAGGTGACCCAGCCCTCAGTGTCTCCCGCTCCAG 2331
QY 1189 GAGCTTATCTCGCCCCATCTCCCAATTAAGTGGGCCCTTGTGCTGTAGGAAGACCAAA 1248
|||||
Db 2332 GAGCTTATCTCGCCCCATCTCCCAATTAAGTGGGCCCTTGTGCTGTAGGAAGACCAAA 2391
QY 1249 GCTCAGGGAAGATAGAGATATGAGATGGAGGGGGAGGACAAAGGGGACAGAGTAGG 1308
Db 2392 GCTCAGGGAAGATAGAGATATGAGATGGAGGGGGAGGACAAAGGGGACAGAGTAGG 2451
QY 1309 GTCTAGCTGGCTATCTCGGCTTACTAACACCCCTGAGGCAATGCCCTTTCTCCA 1368
|||||
Db 2452 GTCTAGCTGGCTATCTCGGCTTACTAACACCCCTGAGGCAATGCCCTTTCTCCA 2511
QY 1369 GCACAAAGCATTGGGGACCTGGAAATATTGGTTCCAGGCTCCTGTCTCTGGACTT 1428
|||||
Db 2512 GCACAAAGCATTGGGGACCTGGAAATATTGGTTCCAGGCTCCTGTCTCTGGACTT 2571
QY 1429 CAGATCTGGGGAGCCCTCCCGCCCTGAAATCCCTGGCTTAGCTTACCTTCTGCTGT 1488
Db 2572 CAGATCTGGGGAGCCCTCCCGCCCTGAAATCCCTGGCTTAGCTTACCTTCTGCTGT 2631
QY 1489 GCACCTAAAAACCTCAGGTCAGAACTAGCAAAAGAGTTTGTATTTTGTGAATG 1548
|||||
Db 2632 GCACCTAAAAACCTCAGGTCAGAACTAGCAAAAGAGTTTGTATTTTGTGAATG 2691
QY 1549 GAGTCTCGTCTGTGCGCCAGGCTGAGGTGAGTAGTGAATCTCCGGCTCACTACAACCT 1608
|||||
Db 2692 GAGTCTCGTCTGTGCGCCAGGCTGAGGTGAGTAGTGAATCTCCGGCTCACTACAACCT 2751
QY 1609 CCACCTCCCTGGGGCTCAAGCGATCTCCACCTCAGCCCGGAGTAGTGGGACTATAG 1668
Db 2752 CCACCTCCCTGGGGCTCAAGCGATCTCCACCTCAGCCCGGAGTAGTGGGACTATAG 2811
QY 1669 GTGTGTACCATCACACCTGGCTAAATTTTGTATTTTGTAGACACAGGGTTTCGCCATG 1728
|||||
Db 2812 GTGTGTACCATCACACCTGGCTAAATTTTGTATTTTGTAGACACAGGGTTTCGCCATG 2871

QY 1729 TTGCCCAAGCTGTGCTTGAATTCCTGAGCTCAAGCAACCTGCCGGCTCGGCCTCCCAA 1788
|||||
Db 2872 TTGCCCAAGCTGTGCTTGAATTCCTGAGCTCAAGCAACCTGCCGGCTCGGCCTCCCAA 2931
QY 1789 GTACTGGGATTTACACGCAGAAAGCCATGCCCAGGCTAGATGTCTTATCCCAATCCT 1848
|||||
Db 2932 GTACTGGGATTTACACGCAGAAAGCCATGCCCAGGCTAGATGTCTTATCCCAATCCT 2991
QY 1849 TTGGCAGCATGACAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTGTAGCCCTCCT 1908
|||||
Db 2992 TTGGCAGCATGACAGCTCCACAGGCGATTTCTTCAAGCAGCTGAAGTGTGTAGCCCTCCT 3051
QY 1909 GGGTTAAAGCCAGCATFAAGAGAAATCCCTTTCTTCTAGGTTTGGAAATGTGTGTG 1962
|||||
Db 3052 GGGTTAAAGCCAGCATFAAGAGAAATCCCTTTCTTCTAGGTTTGGAAATGTGTGTG 3105

RESULT 4
AX236082 1116 bp DNA linear PAT 26-SEP-2001
LOCUS Sequence 13 from Patent WO0164887.
DEFINITION AX236082
ACCESSION AX236082
VERSION AX236082.1 GI:15795889
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1116)
AUTHORS Meyers, R.A.
TITLE Patent: WO 0164887-A 13 07-SEP-2001;
JOURNAL Millennium Pharmaceuticals, Inc. (US)
FEATURES
source location/Qualifiers
1..1116
/organism="Homo sapiens"
/db_xref="taxon:9606"
124..699
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC88255.1"
/db_xref="GI:15795890"
translation="MLGDTGVGKTCFLIQFDGAFLSGTFIATYIDIPRNVTVDG
VRVKLIWDTAGQERFSVTHAYYRDAQLLLYDITNKSFDNIRAMLEIHEYAQR
DVVIMLGNKADMSSEVRISDEGETFLAREYGVPELTSAKTGMNVELAFLAIKELK
YRAGHOADEPSFIIRDYVESQKRSQCSFM"
BASE COUNT 250 a 313 c 322 g 221 t
ORIGIN

Query Match 40.6%; Score 1086; DB 6; Length 1116;
Best Local Similarity 99.5%; Pred. No. 5,1e-280;
Matches 1089; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 37 GGGACATGACGGGACGCCAGCGCGTTCACCCGGGATGCGAGGCCGCCGAGCGCT 96
|||
Db 23 GCGTCGGGACGGGACGCCAGCGCGTTCACCCCGGATGCGAGGCCGCCGAGCGCT 82
|||
QY 97 CCGCGCTTCAGTCCGAGCTCAGGCTCAGGGGCAAGGTGATGCTTCGGGAGACACAG 156
|||||
Db 83 CCGCGCTTCAGTCCGAGCTCAGGCTCAGGGGCAAGGTGATGCTTCGGGAGACACAG 142
|||||
QY 157 GGTTCGGCAAAACATGTTCTCTGATCCAATTCAAAGACGGGGCTTCCTGTCTCCGGAACCT 216
|||||
Db 143 GGTTCGGCAAAACATGTTCTCTGATCCAATTCAAAGACGGGGCTTCCTGTCTCCGGAACCT 202
QY 217 TCATAGCCACCGTCGCGATAGACTTCAGGAACAAGGTGTGACTGTGGATGGCGTGAGAG 276
|||||
Db 203 TCATAGCCACCGTCGCGATAGACTTCAGGAACAAGGTGTGACTGTGGATGGCGTGAGAG 262
QY 277 TGAAGTGCAGATCTGGGACACCGCTGGGAGGAAGGTTCCGAAGGTCACCCATGCTT 336
|||||
Db 263 TGAAGTGCAGATCTGGGACACCGCTGGGAGGAAGGTTCCGAAGGTCACCCATGCTT 322


```
QY 2095 AAACAGATACCGGCGCTAAACTGCTTTTACCTCCCTCCCTACTGAGTCAGGTTAGTGCT 2154
Db 1112 AAACAGATACCGGCGCTAAACTGCTTTTACCTCCCTCCCTACTGAGTCAGGTTAGTGCT 1171
QY 2155 GGGAGGTACCCATTCCTCCGAGTTAAACCAATGCAATATAGTAGTAAACAAGTCATGTGG 2214
Db 1172 GGGAGGTACCCATTCCTCCGAGTTAAACCAATGCAATATAGTAGTAAACAAGTCATGTGG 1231
QY 2215 TATGCTGGGGTAGAGAGAGGGGTAGCAAGTTTCATGTGTCCTTGGTTCACATATCTCC 2274
Db 1232 TATGCTGGGGTAGAGAGAGGGGTAGCAAGTTTCATGTGTCCTTGGTTCACATATCTCC 1291
QY 2275 CAAAGCTCGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGTGCATGACCTGCAGGC 2334
Db 1292 CAAAGCTCGATCCCTGCCATGGGAAGTGGACAGGAAACATGAGTGCATGACCTGCAGGC 1351
QY 2335 ATCTTTACTGACGCTCTGCGGCCCTGGAGGGGAGAGGGGAGGAAGTATGCGCTGC 2394
Db 1352 ATCTTTACTGACGCTCTGCGGCCCTGGAGGGGAGAGGGGAGGAAGTATGCGCTGC 1411
QY 2395 ACATTTCTGAGGCTACTGCATTTGCTTTCAAGSCAGAAATCTTGCTCTGAGCAGTCAGG 2454
Db 1412 ACATTTCTGAGGCTACTGCATTTGCTTTCAAGSCAGAAATCTTGCTCTGAGCAGTCAGG 1471
QY 2455 GCTCCAGTTTGGGCGCGATAAGGAAGTTCTCCGTGGCCCTCCCTCAGGCAGAGCAGGAGG 2514
Db 1472 GCTCCAGTTTGGGCGCGATAAGGAAGTTCTCCGTGGCCCTCCCTCAGGCAGAGCAGGAGG 1531
QY 2515 AGGCTGACATTCGCACTCTCTTCTGGGCGCCAAAGCAGGTTGCGAGAGATCCAAATCCCAT 2574
Db 1532 AGGCTGACATTCGCACTCTCTTCTGGGCGCCAAAGCAGGTTGCGAGAGATCCAAATCCCAT 1591
QY 2575 AGACAGCTCTGGGCTCTTGCAATTTGAGTTTTCAGAATTAACCTGCACTTTTGGAAA 2634
Db 1592 AGACAGCTCTGGGCTCTTGCAATTTGAGTTTTCAGAATTAACCTGCACTTTTGGAAA 1651
QY 2635 GCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2674
Db 1652 GCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1691
RESULT 6
AX236084
LOCUS AX236084 576 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 15 from Patent WO0164887.
ACCESSION AX236084
VERSION AX236084.1 GI:15795891
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 576)
AUTHORS Meyers,R.A.
TITLE 32705, 23224, 27423, 32700, 32712, novel human g-proteins
JOURNAL Patent: WO 0164887-A 15 07-SEP-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
Location/Qualifiers
source 1..576
/db_xref="taxon:9606"
BASE COUNT 141 a 148 c 167 g 120 t
ORIGIN
Query Match 21.5%; Score 576; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.7e-143;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 138 ATGCTCTCTGGGAGACACAGCGCTCGCAAAACATGTTTCTGATCCCAATTCAAAGACGG 197
Db 1 ATGCTCTCTGGGAGACACAGCGCTCGCAAAACATGTTTCTGATCCCAATTCAAAGACGG 60
```

```
QY 198 GCCTTCTCTCGGAAACCTTTCATAGCCACCGTCGGCATAGACTTCAGGAACAGAGTGGTG 257
Db 61 GCCTTCTCTCGGAAACCTTTCATAGCCACCGTCGGCATAGACTTCAGGAACAGAGTGGTG 120
QY 258 ACTGTGGATGGCGTGAAGCTGACAGTCTGGGACACCGCTGGGAGGAAACGGTTC 317
Db 121 ACTGTGGATGGCGTGAAGCTGACAGTCTGGGACACCGCTGGGAGGAAACGGTTC 180
QY 318 CCAAGCGTCACCATGCTTTATTTACAGAGATGCTCAGGCGCTTGCTTCTGCTGTATGACATC 377
Db 181 CCAAGCGTCACCATGCTTTATTTACAGAGATGCTCAGGCGCTTGCTTCTGCTGTATGACATC 240
QY 378 ACCAAATCTCTTTTCGACAACTTCAGGCGCTGGCTCAGTGAATTCATGATGATGCC 437
Db 241 ACCAAATCTCTTTTCGACAACTTCAGGCGCTGGCTCAGTGAATTCATGATGATGCC 300
QY 438 CAGAGGAGCTGGTGATCATGCTGTAGGCAACAGGCGGATATGAGCAGCGAAAGATG 497
Db 301 CAGAGGAGCTGGTGATCATGCTGTAGGCAACAGGCGGATATGAGCAGCGAAAGATG 360
QY 498 ATCCGTTCCGAAGAGGAGAGACCTTTGGCCAGGAGTACGGTCTTCCCTTCTGGAGACC 557
Db 361 ATCCGTTCCGAAGAGGAGAGACCTTTGGCCAGGAGTACGGTCTTCCCTTCTGGAGACC 420
QY 558 AGCGCAAGCAGCTCCAGCTGCTGCTTTCAGGCTTTCTGGCCATCGCCAGGAACCTGAAA 617
Db 421 AGCGCAAGCAGCTGCTGCTGCTTTCAGGCTTTCTGGCCATCGCCAGGAACCTGAAA 480
QY 618 TACCGGCGCGGATCAGGCGGATGAGCCAGCTTCCAGATCCGAGCATATGTAGATGCC 677
Db 481 TACCGGCGCGGATCAGGCGGATGAGCCAGCTTCCAGATCCGAGCATATGTAGATGCC 540
QY 678 CAGAAAGCGCTCCAGCTGCTGCTTTCATGTGA 713
Db 541 CAGAAAGCGCTCCAGCTGCTGCTTTCATGTGA 576
RESULT 7
AF233582 690 bp mRNA linear ROD 02-MAY-2000
LOCUS Mus musculus GTPase Rab37 (Rab37) mRNA, complete cds.
DEFINITION AF233582
ACCESSION AF233582
VERSION AF233582.1 GI:7677421
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 690)
AUTHORS Yu,S., Bennett,M.K., Pavan,D.G. and Scheller,R.H.
TITLE Rab37 is a novel mast cell specific GTPase localized to secretory
granules
JOURNAL FEBS Lett. 470 (1), 61-64 (2000)
MEDLINE 20189834
PUBMED 10728846
REFERENCE 2 (bases 1 to 690)
AUTHORS Luo,Y., Huang,B.C.B., Yu,S., Shen,M. and Masuda,E.S.
TITLE Direct Submission
JOURNAL Submitted (11-FEB-2000) Cell Biology, Rigel, Inc., 240 East Grand
Avenue, South San Francisco, CA 94080, USA
FEATURES
Location/Qualifiers
source 1..690
/organism="Mus musculus"
/db_xref="taxon:10090"
gene 1..690
CDS 1..672
/gene="Rab37"
/codon_start=1
/product="GTPase Rab37"
/protein_id="AAF67162.1"
/db_xref="GI:7677422"
```


/translation="MTGTPGAATAGDGEAPSPFPSPNYDLTKVLLGDSGVKTC
FLIQKDFLSTGIATVGRNKRKVTVDGAKVLQIWDTAGDGRKSVTHAYIRD
AQAALLLYITNQSSFDNRRAWLTTHIYEAQRDVVIMLGNKRDYSSERIVRSEDGET
LAREYGVPMETSAKTMNVELAFLAIKELKYRAGROPDEPSFQIRDYVESQKKRSS
CCSFV"
BASE COUNT 160 a 181 c 201 g 148 t
ORIGIN

Query Match 21.1%; Score 563.6; DB 10; Length 690;
Best Local Similarity 88.6%; Pred. No. 6e-140;
Matches 611; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
QY 42 ATGACGGGACGCCAGCGCGCTGTGCACCCGGGATGGCGAGGCCGCCGAGCGCTCCCG 101
Db 1 ATGACTGGCACACGAGGAGCTGTACCGTGGGGATGGCGAGGCCGCCGAGCGCTCCCG 60
QY 102 CCTGTCAGTCCGAGCTACGACCTACGGGCAAGTGTATGCTTCTGGGAGACACAGCGCTC 161
Db 61 CCTTACGCCGCACTACGATCTACCGGCAAGTGTATGCTTCTGGGAGACTCGGGCGTC 120
QY 162 GGCAAAACATGTTTCTGATCCAAATCAAGACGGGGCTTCTGTCGGGAACCTTCATA 221
Db 121 GGCAAAACCTGTTTCTGATCCAAATCAAGACGGGGCTTCTGTCGGGAACCTTCATA 180
QY 222 GCCACCGTCGCGATAGACTTCAGGAACAAGTGGTGTGATGGATGGCTGAGAGTGAAG 281
Db 181 GCCACCGTCGCGATAGACTTCAGGAATAAGTGGTGTGATGGATGGCTGAGAGTGAAG 240
QY 282 CTGACAGATCTGGGACACGCTGGGACGAACGGTTCGGAAGCGTCAACCATGCTTATTAC 341
Db 241 CTTGAGATCTGGGACATCTGAGGACAGGCGCTTCCGAGTGTGACCATGCTTATTAC 300
QY 342 AGAGATGCTCAGGCTTCTGCTCTGTATGATCATCAACCAAAATCTTCTTCGACAAC 401
Db 301 CGAGATGCTCAGGCTTCTGCTCTGTATGATCATCAACCAACAGTCTCTTTTGACAAC 360
QY 402 ATCAGGCGCTGGCTACATGAGATTCATGATATGCCAGAGGACGTGGTGTATCATCTG 461
Db 361 ATCAGGCGCTGGCTACAGAGATTCATGATATGCCAGAGGACGTGGTGTATCATCTG 420
QY 462 CTAGGCAACAGGCGGATATGAGCAGCGAAGAGTGTATCGTTCCGGAAGACGAGAGACC 521
Db 421 CTAGGCAACAGGCGGATATGAGCAGCGAAGAGTGTATCGTTCTGAAGATGGAGACA 480
QY 522 TTGCGCAGGAGTACGGTGTTCCTTCTGAGACACGCGCAAGACTGCGATGAATGTG 581
Db 481 CTGCGCAGGGAATATGGTGTTCCTTCTGAGACACGAGTCCCAAGACTGGCATGAACGTG 540
QY 582 GAGTTAGCTTCTGCGCATCGCAAGAACTGAATACCGGCGCGGCATCAGCGCGAT 641
Db 541 GAGTTGCGCTTCTGCGCAATGCGCAAGAACTGAATACCGGCGCGGCATCAGCGCGAT 600
QY 642 GAGCGCAGCTTCCAGATCCGAGACTATGATAGTCCCAAGAAAGCGCTCCAGCTGCTGC 701
Db 601 GAGCGCAGCTTCCAGATCCGAGACTATGATAGTCCCAAGAAAGCGCTCCAGCTGCTGC 660
QY 702 TCCTTCATGTGAATCCAGGGGGCAGAGAG 731
Db 661 TCCTTTGTGTACCCCTAGGGGCTAAGAG 690

RESULT 8
BC007681
LOCUS BC007681 1513 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, RAB26, member RAS oncogene family, clone MGC:3503
IMAGE:3627067, mRNA, complete cds.
ACCESSION BC007681
VERSION BC007681.1 GI:14043378
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1513)
Strausberg,R.
Direct Submission
Submitted (11-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nih.gov
Shevchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R.,
Lim,M., Maduro,Q.L., Masfello,C., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Snyder,B., Stantripop,S., Thomas,P.J.,
Tongson,E.E., Touchman,J.W., Tsugeon,C., Vogt,J.L., Walker,M.A.,
Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 12 Row: m Column: 2
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5931611.

FEATURES
Location/Qualifiers
1..1513
/organism="Homo sapiens"
/db_xref="LocusID:25837"
/db_xref="taxon:9606"
/clone="MGC:3503 IMAGE:3627067"
/tissue_type="Uterus, endometrium adenocarcinoma"
/clone_lib="NIH_MGC_44"
/lab_host="DH10B-R"
/note="Vector: pOTR7"
187..759
/codon_start=1
/product="RAB26, member RAS oncogene family"
/protein_id="AAH07681.1"
/db_xref="GI:14043379"
/translation="MLVGSVGVKTCCLLVRFKDFLAGTFFSTVIGIDFRNKVLVDVG
VKVQLQMDTAGOERFSVTHAYIRDAHALLLLYDVTNRKASFDNIQAWLTETHEYAOH
DVALMLGNKVDSAHERVVKREDGEKLAKEYGLPFMETSAKTGLNVLDAFTAKELK
QSRMKAPSEPRFLHDYVREGRGASCCRP"
BASE COUNT 326 a 452 g 449 g 286 t
ORIGIN

Query Match 11.8%; Score 315.6; DB 9; Length 1513;
Best Local Similarity 71.2%; Pred. No. 2.4e-73;
Matches 417; Conservative 0; Mismatches 169; Indels 0; Gaps 0;
QY 116 CTACGACCTCAGCGGCAAGGTGATGCTTCTGGGAGACACAGCGCTCGGCAAAACATGTTT 175
Db 165 CTACGAGCTCGCTTCAAGGTCATGCTGGTGGGGGACCTCGGGTGTGGGGAAGACCTGCT 224
QY 176 CCTGATCAATTCAAAGACGGGGCTTCTTCCGGAACCTTCATACGCCACCGTCGGCAT 235
Db 225 GCTGGTGCATCAAGGATGGTCTTCTTCCGCGGGGACCTTCATCCTCCACCGTAGGCAT 284
QY 236 AGACTTCAGGAACAAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 295
Db 285 TGACTTCGGAACAAGATTTCTGGACGTGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 344
QY 296 CACCGCTGGCAGGAACGGTTCGGAAGCGTCAACCATGCTTATPACAGAGATGCTCAGCG 355
Db 345 CACAGCTGGTCAAGGAGGGTTCGCGAGTGTACCCAGTGTACCCAGTGTACCCAGTGTCTATCG 404

CDS		/gene="v46133" 4..576 /gene="v46133" /note="RAS-related protein" /codon_start=1 /product="RAB-26" /protein_id="BA084707.1" /db_xref="GI:5931612" VKVQLQMDTGAQGERFRSVHAYYRDHAHLLLDVYTNKASPDNIQAWLTFEIEYAAQH DVALMLGNKVDSAHERRVKREDEKLAKEYGLPFMETSAKTGLNVDLAFATAKELK QRSMKAPSEPRFLHDYVKREGGASCRP"		BASE COUNT 273 a 379 c 395 g 273 t	
ORIGIN		11.4%; Score 304.6; DB 9; Length 1320; Query Match 71.1%; Pred. No. 2e-70; Best Local Similarity 0; Mismatches 164; Indels 0; Gaps 0; Matches 403; Conservative			
Qy	135	GTGATGCTTCTGGGACACAGCGCTCGGCAAAACATGTTCTCTGATCCAAATTCAAAGAC	194		
Db	1	GTCATGCTGTGGGGACTCGGGTGTGGGAAGACCTGCTGCTGGTCCGATTCAAGGAT	60		
Qy	195	GGGCGCTTCCTGTCGGGAACCTTCATAGCACCGCTGGCGATAGACTTCAGGAACAAGGTG	254		
Db	61	GGTGCTTCTCGGGGGACCTTCATCTCCACCGTAGGCATTGACTTCGGGAACAAGATT	120		
Qy	255	GTGACTGTGATGGCTGAGAGTGAGCTGCAGATCTGGGACACCGCTGGCGAGGAACGG	314		
Db	121	CTGACGTGATGGTGTGAAGGTGAAGCTGCAGATGTGGGACACAGCTGGTCAGAGCGG	180		
Qy	315	TTCCGAAGCGCTACCCATGCTATTACAGAGATGCTCAGCGCTTCTGCTGTGATGAC	374		
Db	181	TTCCGCAGTGTACCATGCTACTACCGGATGCTCATGCTCTGCTGCTCTACGAT	240		
Qy	375	ATCACCAACAATTTCTTTTCGACACATCAGGGGCTGGCTCACTGAGATTCAATGAGTAT	434		
Db	241	GTCAACCAACAAGGCTCTTTGACACATCCAGGCGCTGGCTCAGCGAGATCCACGATAC	300		
Qy	435	GCCAGAGGACCTGTGTATGCTGCTAGCGACACAAAGCGGATATGAGCAGCAAGA	494		
Db	301	GCCAGCAGCAGCTGGCGCTTCATCTGCTGGGGAACAAGTGGATCTGCCCATGAGCGT	360		
Qy	495	GTGATCGCTTCCGAAGACGAGAGACCTTGGCCAGGGAGTAGCGGTTCCTTCCTGGAG	554		
Db	361	GTGTTGAAGAGGAGGAGCGGAGAGCTGGCCAAAGGATATGAGTGCCTTTCATGAG	420		
Qy	555	ACAGCGCCAAAGACTGGCATGATGTGGAGTTAGCCCTTCTGGCCATGCCAAGAACTG	614		
Db	421	ACCAGCGCAAGACGGGCGCTTCAACGTGGAGTTGGCCCTTCCAGCGCATAGCAAGAGT	480		
Qy	615	AAATACCGCGCGGCGATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAG	674		
Db	481	AAGCAGCGCTCCATGAAGGCTCCACGAGCGCGCTTCCGGCTGCATGATTACGTTAAG	540		
Qy	675	TCCCAGAAGAAGCGCTCCAGCTGCTGC	701		
Db	541	AGGAGGTCGAGGGGCTCTCTGCTGC	567		
RESULT 11		573 bp mRNA linear PRI 01-MAY-2002			
AF498952		AF498952			
LOCUS		AF498952.1 GI:20379079			
DEFINITION		Homo sapiens small GTP binding protein RAB26 (RAB26) mRNA, complete cds.			
ACCESSION		AF498952			
VERSION		AF498952.1			
KEYWORDS		GI:20379079			
SOURCE		Homo sapiens.			
ORGANISM		Homo sapiens.			
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 573)			

RESULT 12	AR062279	1340 bp	DNA	linear	PAT 29-SEP-1999
LOCUS	Sequence 2 from patent US 5843717.				
DEFINITION	AR062279				
ACCESSION	AR062279.1	GI:5989970			
VERSION	Unknown.				
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 1340)				
AUTHORS	Hillman,J.L. and Guegler,K.J.				
TITLE	Rab protein				
JOURNAL	Patent: US 5843717-A 2 01-DEC-1998;				
FEATURES	Location/Qualifiers				
Source	1..1340				
BASE COUNT	276 a 387 c 403 g 274 t				
ORIGIN	/organism="unknown"				
Query Match	10.9%; Score 292.6; DB 6; Length 1340;				
Best Local Similarity	70.9%; Pred. No. 3.4e-67;				
Matches	416; Conservative 0; Mismatches 169; Indels 2; Gaps 2;				
QY	116 CTACGACCTCACGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTT 175				
Db	116 CTACGACCTCACGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTT 175				
QY	176 CTTGA-TCCAAATCAAGACGGGGCTTCTGTCGGGAACCTTCATAGCCACCGTCGGCA 234				
Db	176 CTTGA-TCCAAATCAAGACGGGGCTTCTGTCGGGAACCTTCATAGCCACCGTCGGCA 234				
QY	81 GCTGGTGGCATCAAGGATGCTGCTTCTGTCGGGGACCTTCATCTCCACCGT-AGCA 139				
Db	81 GCTGGTGGCATCAAGGATGCTGCTTCTGTCGGGGACCTTCATCTCCACCGT-AGCA 139				
QY	235 TAGACTTCAGGAACAAGGTGCTGACTGTGGATGCTGGAGTGAAGCTGCACATCTGGG 294				
Db	235 TAGACTTCAGGAACAAGGTGCTGACTGTGGATGCTGGAGTGAAGCTGCACATCTGGG 294				
QY	140 TTGACTTCGGACAAAGTCTCGAGCTGGATGCTGTGAAGGTGAAGCTGCAGATGTGGG 199				
Db	140 TTGACTTCGGACAAAGTCTCGAGCTGGATGCTGTGAAGGTGAAGCTGCAGATGTGGG 199				
QY	295 ACACCGCTGGGACGAGACGGTTCGGAAGCGCTACCCATGCTTATTACAGAGATGCTCAGG 354				
Db	295 ACACCGCTGGGACGAGACGGTTCGGAAGCGCTACCCATGCTTATTACAGAGATGCTCAGG 354				
QY	200 ACACAGCTGCTCAGGACGGTTCGCGAGTGTACCCATGCTTACTACCGGATGCTCATG 259				
Db	200 ACACAGCTGCTCAGGACGGTTCGCGAGTGTACCCATGCTTACTACCGGATGCTCATG 259				
QY	355 CTTGCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 414				
Db	355 CTTGCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 414				
QY	260 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 319				
Db	260 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 319				
QY	415 TCACTGAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 474				
Db	415 TCACTGAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 474				
QY	320 TGACCGAGATCCAGAGTACGCGCCAGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 379				
Db	320 TGACCGAGATCCAGAGTACGCGCCAGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 379				
QY	475 CGGATATGACGACGAAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 534				
Db	475 CGGATATGACGACGAAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 534				
QY	535 ACAGTGTTCCTTCTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 594				
Db	535 ACAGTGTTCCTTCTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 594				
QY	440 ATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 499				
Db	440 ATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 499				
QY	595 TGCCATCGCAAGGAACCTGAATACCGCGCGCGGATGATGATGATGATGATGATGATGATG 654				
Db	595 TGCCATCGCAAGGAACCTGAATACCGCGCGCGGATGATGATGATGATGATGATGATGATG 654				
QY	500 CAGCCATAGCAAGGAGTGTGAACAGCGCTTCCATGAAGGCTCCACGAGCGCGGCTTCC 559				
Db	500 CAGCCATAGCAAGGAGTGTGAACAGCGCTTCCATGAAGGCTCCACGAGCGCGGCTTCC 559				
QY	655 AGATCGGAGCTATGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 701				
Db	655 AGATCGGAGCTATGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 701				
QY	560 GGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 606				
Db	560 GGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 606				
RESULT 13	AY061826	2536 bp	mRNA	linear	INV 10-NOV-2001
LOCUS	Drosophila melanogaster GH21984				
DEFINITION	full length cDNA.				
ACCESSION	AY061826				
VERSION	AY061826.1	GI:16902019			
KEYWORDS	FLI_CDNA.				

SOURCE	Drosophila melanogaster.
ORGANISM	Drosophila melanogaster
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
AUTHORS	Stapleton,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George,R., Gonzalez,M., Guarin,H., Li,P., Liao,G., Miranda,A., Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Rubin,G.M. and Celnikier,S.
TITLE	Direct Submission
JOURNAL	Submitted (02-NOV-2001) Berkeley Drosophila Genome Project, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA
COMMENT	Sequence submitted by: Berkeley Drosophila Genome Project Lawrence Berkeley National Laboratory Berkeley, CA 94720 This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our Web site (http://fruitfly.berkeley.edu) or send email to cdna@fruitfly.berkeley.edu.
FEATURES	Location/Qualifiers
source	1..2536
gene	/organism="Drosophila melanogaster" /strain="y; cn bw sp" /db_xref="taxon:7227" /map="78C9-78D1" 1..2536
CDS	/gene="CG7605" /note="alignment with genomic scaffold AE003594" /db_xref="FLYBASE:FBgn0037072" 1172..2338 /note="Longest ORF" /product="GH21984p" /protein_id="AAL27637.1" /db_xref="GI:16902020" /db_xref="FLYBASE:FBgn0037072" /translation="MASTAVLGSGEGDFGAGGPPAGSHPDASMSDDVDFDAETT QARIEELRRPFGDSYNPAPASVSASITTTTQOOOHHNPSSHHSSHHOPSHH HHHHHSSLSLTGSHYHDDAIPAVQSRATGYPGRPSREAMQMTAYGTDDYDYN DGRWSYRDEVDMPAPNAHQPFDDTVNHKTILIGDSGVGKTSFLVYNTGEFLRG SFSATVGIALTNKVVVVDGTRVKIOWDAGQSRFSRTHAYVRDAHALLLVDVTK TYYDNIRALGELREYAEQDVIVLIGNKADCSEKROVKRDEGERLGRHNHVPMET SAKTGLNVELSFTAVARQLKSRGHEGDDGRFVHDFVNDTKARSVCAQCRNN"
BASE COUNT	831 a 621 c 624 g 460 t
ORIGIN	
Query Match	8.8%; Score 234.2; DB 3; Length 2536;
Best Local Similarity	65.7%; Pred. No. 1.8e-51;
Matches	357; Conservative 0; Mismatches 183; Indels 3; Gaps 1;
QY	131 CAAGGTGATGCTTCTGGGAGACACAGGCTCGGCAAAACATGTTCTGTGATCCAAATCAA 190
Db	131 CAAGGTGATGCTTCTGGGAGACACAGGCTCGGCAAAACATGTTCTGTGATCCAAATCAA 190
QY	1744 CAGACCATCTCTCTGGGATTCGGCGCTGGGAAACACCTCTCTCTGTGTCATACAA 1803
Db	1744 CAGACCATCTCTCTGGGATTCGGCGCTGGGAAACACCTCTCTCTGTGTCATACAA 1803
QY	191 AGACGGGCGCTTCTCTGCGGAACCTTTCATAGCCACCGTCGCGATAGACTTCAGGAACAA 250
Db	191 AGACGGGCGCTTCTCTGCGGAACCTTTCATAGCCACCGTCGCGATAGACTTCAGGAACAA 250
QY	1804 TACGGGCGAGTTCGCGACTGGGTTCTCTTTCCCGCACAGTGGGATTCGCGTTACGAACAA 1863
Db	1804 TACGGGCGAGTTCGCGACTGGGTTCTCTTTCCCGCACAGTGGGATTCGCGTTACGAACAA 1863

QY	251	GGTGGTGAAGTGGCGGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGCGCAGGA	310
Db	1864	AGTGGTGGTCTCGATGGAACGCGCGTCAAGCTGCAAAATCTGGGACACAGCTGCTCAGGA	1923
QY	311	ACGTTTCGGAAGCGTCAACCATGCTTATTACAGAGATGCTCAGGCCCTGCTTCTGCTGTA	370
Db	1924	CGAATTCGGAGCGTTATCCACGCGCTATTATCGGACGCGCACGCTCTACTGCTGCTGTA	1983
QY	371	TGACATCACCACAAATCTCTTCGACACATCAGGCGCTGGCTCACTGAGATTCATGA	430
Db	1984	CGAGTGAACCAAGACACCTATGACAACTTCGCGCTGGCTGGCGGAGATCCCGGA	2043
QY	431	GTATGCCAGAGACGCTGTGATCTGCTAGGCAACAAGCGGATATGA---CGAC	487
Db	2044	GTACGCGCAGGAGCGTGTGATCTGTTAATAGGCAACAAGCGCGCTGCTGCGGCGAG	2103
QY	488	CGAAGAGTGTATCGTTCGGAAGACGAGAGACCTTGGCGAGGAGTACGGTGTTCCTT	547
Db	2104	CGAGCGCAGGTGAAGCGGGAGGATGGGAGCGTTTGGGCGGAGCACACGTCGCCCTT	2163
QY	548	CCTGGAGACACCGCCGAGCTGCATGATGTGAGTTCAGCTTTCTGGCCATCGCCAA	607
Db	2164	CATGGAGCTTCGCGCAGGACGAGCTCAATGTGGAGCTGCTTCACAGCGGTGGCCAG	2223
QY	608	GGAATCAATACCGCGCGCGGCATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTA	667
Db	2224	GCAACTAAGAGTTCGCGGCTACGAGCAGCGGATGATGGAAGTTCATGTGCATGATT	2283
QY	668	TGT	670
Db	2284	TGT	2286
RESULT 14			
HS941F9/C			
LOCUS	HS941F9	127587 bp	DNA linear PRI 12-DEC-1999
DEFINITION	Human DNA sequence from clone CTA-941F9 on chromosome 22q13. Contains the 3' end of the FBLN1 gene for Fibulin 1 isoforms B, C and D, the first exon of the gene for a novel protein (the ortholog of mouse brain protein E46), ESTs, STSs, GSSs and two putative CpG islands, complete sequence.		
ACCESSION	295331		
VERSION	295331.2	GI:6572282	
KEYWORDS	HTG; E46; FBLN1; Fibulin 1.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 127587)		
AUTHORS	Connor, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk		
COMMENT	requests: clones@res.sanger.ac.uk On Dec 13, 1999 this sequence version replaced gi:2281936. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submissions corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unseq' feature key. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence		

was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre Chromosome 22 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr22>
CTA-941F9 is from the human BAC library described in U-J. Kim et al. (1996) Genomics 34, 213-218.
VECTOR: pBeloBAC11
This sequence is the entire insert of clone CTA-941F9 The true left end of clone RP3-398C22 is at 87881 in this sequence. The true left end of clone LL22NC03-38E2 is at 9039 in this sequence. The true right end of clone LL22NC03-38E2 is at 54042 in this sequence.

FEATURES	Location/Qualifiers
source	1..127587 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="22" /map="q13"
misc_feature	/clone="CTA-941F9" 18..560 /note="match: GSS: Em:B14132"
misc_feature	25..461 /note="match: GSS: Em:AQ777382" complement(1107..1512) /note="match: GSS: Em:AQ060472"
misc_feature	1436..1467 /note="16 copies 2 mer gt 84 conserved"
repeat_region	2424..2647 /note="7 copies 32 mer 94 conserved"
repeat_region	2664..2735 /note="MLT2B repeat: matches 381..448 of consensus"
repeat_region	2764..3168 /note="MLT2B repeat: matches 1..409 of consensus"
repeat_region	3169..3439 /note="HBRVL repeat: matches 5384..5654 of consensus"
repeat_region	3440..3752 /note="AluX repeat: matches 1..311 of consensus"
repeat_region	3753..9139 /note="HBRVL repeat: matches 1..5384 of consensus"
repeat_region	9140..9577 /note="MLT2B repeat: matches 1..444 of consensus"
repeat_region	9778..9851 /note="MIR repeat: matches 73..141 of consensus"
gene	10369..14852 /gene="FBLN1"
mRNA	<10369..10819 /gene="FBLN1" /product="bk941F9.2 (Fibulin 1 (isoform C))" /notes="match: CDNAS: Em:AF013752 Em:AF051400 Em:X53743 Em:X70853 Em:AL050095 Em:X75285 Em:X82494; match: ESTs: Em:R25317 Em:AA226244 Em:AV064777 Em:R62365 Em:AA508812 Em:AA041958 Em:AA626088 Em:A1092723 Em:A1130996 Em:A1188632 Em:A1273494 Em:H02670 Em:A1893154 Em:A1167845 Em:A1890847 Em:AA368999 Em:R64034 Em:F16897 Em:AA142940 Em:AV064836 Em:R23566 Em:AA010959 Em:R76273 Em:R68191 Em:R68138 Em:C02271 Em:AA55258 Em:A1066730 Em:A1087421 Em:A1493361 Em:A1803546 Em:A1828352 Em:C18390 Em:AA781778 Em:R12400 Em:T49888 Em:A1084677 Em:A1553895 Em:A1858475" /evidence="not_experimental"
CDS	<10369..10723 /gene="FBLN1" /note="match: proteins: Sw:P23144 Sw:Q08878" /codon_start=2 /evidence="not_experimental" /product="bk941F9.2 (Fibulin 1 (isoform C))" /protein_id="CAB62860.1" /db_xref="GI:6572283" /translation="RCERLPCHENRECSKPLRITYYHLSPFTNIQAPVAVFRMGPS AVPGDSNQIAITGNEEGFTTRKVSHPHSGWALTKPVPEPRDLLLTVMKDLRSRGTG SSFVAKLFIIVSAEL" 10374..10803 /gene="FBLN1" /note="match: STS: Em:R76273"
misc_feature	

```
polyA_signal complement(10783..10789)
polyA_site /gene="FBLN1"
polyA_site /gene="FBLN1"
repeat_region 11114..11149
repeat_region /note="MIR repeat: matches 109..144 of consensus"
repeat_region 11527..11572
repeat_region /note="MIR repeat: matches 75..120 of consensus"
repeat_region 11651..11934
repeat_region /note="L2 repeat: matches 1992..2328 of consensus"
repeat_region 12001..12270
repeat_region /note="L2 repeat: matches 2454..2750 of consensus"
misc_feature 12337..13158
misc_feature /gene="FBLN1"
misc_feature /note="match: STS: Em:X53742"
misc_feature <12345..13158
misc_feature /gene="FBLN1"
misc_feature /product="bk941f9.2 (Fibulin 1 (isoform B))"
misc_feature /note="match: CDNAS: Em:AB007943 Em:AF078528 Em:Y14610
misc_feature Em:AF011357 Em:X53742 Em:AF019887; match: STS: Em:R78704
misc_feature Em:R33281 Em:R32733 Em:H01971 Em:H02060"
misc_feature /evidence=not_experimental
misc_feature <12345..12449
misc_feature /gene="FBLN1"
misc_feature /note="match: proteins: Sw:P23143"
misc_feature /codon_start=1
misc_feature /evidence=not_experimental
misc_feature /product="bk941f9.2 (Fibulin 1 (isoform B))"
misc_feature /protein_id="CAB62961.1"
misc_feature /db_xref="GI:6572284"
misc_feature /translation="KSKKGRQNTPAGSSKEDCRVLPWKQGLETHLDA"
misc_feature 12449..12703
misc_feature /gene="FBLN1"
misc_feature /note="match: STS: Em:G59664"
misc_feature 12450..12849
misc_feature /gene="FBLN1"
misc_feature /note="match: STS: Em:G11168"
misc_feature complement(12733..13139)
misc_feature /note="match: STS: Em:H01971"
misc_feature 12895..13143
misc_feature /note="L2 repeat: matches 2419..2705 of consensus"
misc_feature complement(12974..13280)
misc_feature /note="match: STS: Em:G08093; match: STS: Em:G08093"
polyA_signal 13138..13143
polyA_signal /gene="FBLN1"
polyA_site 13158
polyA_site /gene="FBLN1"
polyA_site 13237..13536
polyA_site /note="AluSp repeat: matches 1..300 of consensus"
repeat_region 13974..14045
repeat_region /note="MIR repeat: matches 50..122 of consensus"
repeat_region 14352..14607
repeat_region /note="MIR repeat: matches 2..262 of consensus"
repeat_region 14791..15080
repeat_region /note="AluSx repeat: matches 1..290 of consensus"
repeat_region 15714..15956
repeat_region /note="MLTIC repeat: matches 300..451 of consensus"
repeat_region 16332..16625
repeat_region /note="MLTIC repeat: matches 1..310 of consensus"
repeat_region 16637..16815
repeat_region /note="MLTIB repeat: matches 1..190 of consensus"
repeat_region 16816..17125
repeat_region /note="AluSg repeat: matches 1..310 of consensus"
repeat_region 17126..17338
repeat_region /note="MLTIB repeat: matches 190..381 of consensus"
repeat_region 17506..17806
repeat_region /note="AluYb8 repeat: matches 1..318 of consensus"
repeat_region 17807..18030
repeat_region /note="7 copies 32 mer 64 conserved"
repeat_region 17808..18031
repeat_region /note="112 copies 2 mer aa 62 conserved"
repeat_region 18046..18303
```

```
repeat_region /note="HAL1 repeat: matches 351..597 of consensus"
repeat_region 18304..18597
repeat_region /note="AluSg repeat: matches 1..309 of consensus"
repeat_region 18598..18822
repeat_region /note="HAL1 repeat: matches 597..825 of consensus"
repeat_region 19070..19305
repeat_region /note="L1MC5 repeat: matches 6484..6730 of consensus"
repeat_region 19306..19621
repeat_region /note="AluSg repeat: matches 1..310 of consensus"
repeat_region 19622..20401
repeat_region /note="L1MC5 repeat: matches 6730..7471 of consensus"
repeat_region 20402..20711
repeat_region /note="AluSx repeat: matches 1..312 of consensus"
repeat_region 20712..20770
repeat_region /note="L1MC5 repeat: matches 7471..7531 of consensus"
repeat_region 20851..21158
repeat_region /note="AluSg repeat: matches 1..301 of consensus"

Query Match 8.0%; Score 214.8; DB 9; Length 127587;
Best Local Similarity 80.3%; Pred. No. 7.5e-46;
Matches 252; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 1526 TTGTGTTTATTTTGGAAATGGAGTCTGTCGCCCCAGGCTGAGGTGCAGTACT 1585
Db 59011 TTTGTTTTCTTTTTTGAACGGAGTCTCACTGTCTACCCAGGCTGGAGTGCATATG 58952
QY 1586 GCAATCTCGGCTCACTACAACCTCCACTCCCTGGGCTCAAGCGATCCCTCCACCTCAGC 1645
Db 58951 GCGATCTCGGCTTATGAGCGCTCTGCCCTCCGGGATCAAGTATCTCTCGCTCAGC 58892
QY 1646 CCCCAGATAGCTGGGACTATAGGTGTGTTACCATCACACTGGCTGAATTTTGTATTTT 1705
Db 58891 CTCCCAAGTAGCTGGAATTACAGGTGCATACCACACACTGGCTGAATTTTATATTTT 58832
QY 1706 TGTACACACAGGTTTCCGCTATGTTGCCAGCTGGTCTTGAATTCCTGAGCTCAGCAA 1765
Db 58831 AGTAGAGATGGGGTTTCAACATGTTGCCAGGCTGGTCTTGAATTCCTGAGCTCAGCAA 58772
QY 1766 CTGCGCGGCTCGGCTCCCAAAGTACTGGGATTTACACGACGAAGCACCATGCCAGCC 1825
Db 58771 TGTGCTGCTCGGCTCCCAAAGTCCCGGATTACAGCAGACGACCATGCCAGCC 58712
QY 1826 TAGATGTGTTCTTAT 1839
Db 58711 TGGTTTCTGTTTT 58698

RESULT 15
AC010352/c
LOCUS AC010352 102332 bp DNA linear PRI 18-APR-2000
DEFINITION Homo sapiens chromosome 5 clone CTD-2022F20, complete sequence.
ACCESSION AC010352
VERSION AC010352.4 GI:7109394
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 102332)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 102332)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 102332)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (29-FEB-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 102332)
```

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (18-APR-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Feb 29, 2000 this sequence version replaced gi:6606056.
Draft Sequence Produced by DOE Joint Genome Institute

www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.4% of Sequence;
Estimated Total Number of Errors is 1.5.
STS Content:
WI-12799 G24393
SHGC-13554 G14788.

FEATURES
Source Location/Qualifiers
1..102332
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2022F20"
BASE COUNT 29578 a 22934 c 23227 g 26593 t
ORIGIN

Query Match 8.0%; Score 214.2; DB 9; Length 102332;
Best Local Similarity 82.6%; Pred. No. 1e-45;
Matches 257; Conservative 0; Mismatches 53; Indels 1; Gaps 1;
QY 1522 GAGTTTGTGTTTATTTTGTAAATGGAGTCCTGTCGCCAGGCTGAGGTGCAG 1581
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 22554 GATTTTATTTTATTTTGTGACACAGTCCTGCTCTGTCGCCAGGCTGGAATGCAG 22495
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
QY 1582 TAGTGAATCTCCGCTCACTACAACTCCACTCCCTCGGGCTCAAGCGATCCTCCACCT 1641
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 22494 TGGTGAATCTCGGCTCCCTGCAACCTCCACTTCCT -GGTTGAAGCGATTCTCTGCT 22436
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
QY 1642 CAGCCCGGAGTAGTGGGACTATAGGTGTGTACCATCACACCTGGCTAAATTTTGTAT 1701
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 22435 CAGCCTCCTGAGTAGTGGGACTACAGGCGTGTGCCACCACACCCGGCTAAATTTTGTAT 22376
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
QY 1702 TTTTGTAGACAGAGGTTTTCGCATGTTGCCAGGCTGGTCTTGAATCCTGAGCTCAA 1761
|| || |||| || |||| || |||| || |||| || |||| || |||| || ||||
Db 22375 TTTTGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAG 22316
|| || |||| || |||| || |||| || |||| || |||| || |||| || ||||
QY 1762 GCAACCTGGCGGCTCGGCTCCCAAAGTACTGGGATTACAGCAGAAGSCACCATGCC 1821
|| || |||| || |||| || |||| || |||| || |||| || |||| || ||||
Db 22315 GCAATCTGCCATCTCGGCTCCCAAAGTCTGGGATTACAAGCATGAGCCACCATGCC 22256
|| || |||| || |||| || |||| || |||| || |||| || |||| || ||||
QY 1822 AGCTAGATCT 1832
|| || || || ||
Db 22255 AGACTATTTAT 22245

Search completed: November 17, 2002, 16:01:28
Job time : 4970 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 17, 2002, 13:08:43 ; Search time 368 Seconds
(without alignments)
16363.705 Million cell updates/sec

Title: US-09-817-199A-1

Perfect score: 2674

Sequence: 1 ttgcctcgccggccgcact.....aaaaaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2581.6	96.5	2623	22	AAS27036
2	2240	83.8	2323	22	AAI60852
3	1286	48.1	1316	22	AAS27458
4	1099.6	41.1	1106	22	AAI59066
5	1079	40.4	1109	22	AAH75184
6	873	32.6	875	22	AAH83696
7	795.4	29.7	797	21	AAA52610
8	568	21.2	576	24	ABA02774
9	550.8	20.6	843	20	AAH86720
					CDNA encoding nove
					Human polynucleoti
					CDNA encoding nove
					Human polynucleoti
					Nucleotide sequenc
					Human Ras protein
					Eosinophil activat
					Human degranulatio
					DNA sequence encod

10	480	18.0	576	24	ABA02773	Mouse degranulatio
11	468.6	17.5	475	23	AAS81431	DNA encoding nove
12	292.6	10.9	1340	19	AAV65197	Human RAB protein,
13	274	10.2	443	21	AAH97117	Mouse Exol04 nucle
14	237	8.9	2001	23	ABL29523	Drosophila melanog
15	209.4	7.8	964	22	AAS27040	CDNA encoding nove
16	207.4	7.8	24000	21	AAA88551	Human dual-specific
17	206.6	7.7	1786	22	AAH08354	Human secreted pro
18	204.8	7.7	10859	22	AAH04720	Human reproductive
19	204.8	7.7	10859	23	ABL97627	Human testicular a
20	203.4	7.6	138169	21	AAA34791	Human adenosine re
21	203.4	7.6	141589	21	AAF20913	Human ELAM-1 polyn
22	203.4	7.6	141589	21	AAF21127	Human low adenosin
23	203.4	7.6	141589	21	AAF21152	Human low adenosin
24	203.4	7.6	141589	21	AAA35005	Human adenosine re
25	203.4	7.6	141589	21	AAA35030	Human adenosine re
26	203.4	7.6	146981	21	AAF21442	Human ELAM-1 polyn
27	203.4	7.6	209273	21	AAF21437	Human factor-relat
c 28	203	7.6	409	22	AAK70203	Human immune/haema
c 29	202.8	7.6	27579	22	AAK66484	Human immune/haema
c 30	202.6	7.6	29222	24	ABL39412	Human electron-tra
c 31	202.4	7.6	10468	22	AAK77961	Human immune/haema
c 32	201.8	7.5	103747	24	ABQ88139	Human osteoblast d
c 33	201.6	7.5	52845	22	AAK71437	Human immune/haema
34	201.4	7.5	4969	22	ABA16169	Human nervous syst
35	201.4	7.5	6720	24	ABO88138	Human osteoblast d
c 36	201.2	7.5	33353	22	AAK70003	Human immune/haema
37	201	7.5	7428	22	ABA07164	Human pancreatic c
38	200.8	7.5	21835	22	AAK66483	Human immune/haema
c 39	200.6	7.5	21833	22	AAL35910	Human musculoskele
c 40	200.4	7.5	10663	22	ABA18961	Human nervous syst
c 41	200.4	7.5	31865	22	AAK85330	Human immune/haema
c 42	200.4	7.5	31865	22	AAK85331	Human immune/haema
c 43	200.2	7.5	21777	22	AAK74096	Human immune/haema
c 44	199.8	7.5	1130	22	ABA20629	Human nervous syst
c 45	199.8	7.5	1130	22	AAL37008	Human musculoskele

ALIGNMENTS

RESULT 1

AAS27036

ID AAS27036 standard; cDNA; 2623 BP.

XX AC AAS27036;

XX DT 07-NOV-2001 (first entry)

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX DE CDNA encoding novel signal transduction pathway protein, Seq ID 71.

XX The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
CC pathway protein coding sequences and PCR primers of the invention.
XX

Query Match 96.5%; Score 2581.6; DB 22; Length 2623;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2606; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY	42	ATGACGGGCACGCCAGGCGCGTGGCCACCGGATGGCGAGCGCCCGGAGCGCTCCCGG	101
DB	12	ATGACGGGCACGCCAGGCGCGTGGCCACCGGATGGCGAGCGCCCGGAGCGCTCCCGG	71
QY	102	CCCTGCAGTCCGAGCTACGACCTCAGCGGCAAGGTGATGCTTCTGGGACACACAGCGT	161
DB	72	CCCTGCAGTCCGAGCTACGACCTCAGCGGCAAGGTGATGCTTCTGGGACACACAGCGT	131
QY	162	GGCAAAACATGTTCTGTATCAATTCAAAGACGGGGCTTCTGTCGGAAACCTTCAT	221
DB	132	GGCAAAACATGTTCTGTATCAATTCAAAGACGGGGCTTCTGTCGGAAACCTTCAT	191
QY	222	GCCACGCTCGGCATAGACTTCAGAACAGGTGGTGACTGTGGATGGCGTGAGATGA	281
DB	192	GCCACGCTCGGCATAGACTTCAGAACAGGTGGTGACTGTGGATGGCGTGAGATGA	251
QY	282	CTGCAGATCTGGACACCGCTGGGAGGAACGTTCCGAAGCGTCAACCATGCTATTAC	341
DB	252	CTGCAGATCTGGACACCGCTGGGAGGAACGTTCCGAAGCGTCAACCATGCTATTAC	311
QY	342	AGAGATGCTCAGGCGCTTGTCTGTGTATGACATCACCAACAAATCTTCTTCGACAC	401
DB	312	AGAGATGCTCAGGCGCTTGTCTGTGTATGACATCACCAACAAATCTTCTTCGACAC	371
QY	402	ATCAGGCGCTGCTCACTAGATTATAGTATGCCACAGGAGCGTGTGATCATGCTG	461
DB	372	ATCAGGCGCTGCTCACTAGATTATAGTATGCCACAGGAGCGTGTGATCATGCTG	431
QY	462	CTAGGCAACAAGCGGATATGACAGCGAAAGTATCGCTTCCGAAGCGGAGAGACC	521
DB	432	CTAGGCAACAAGCGGATATGACAGCGAAAGTATCGCTTCCGAAGCGGAGAGACC	491
QY	522	TTGGCCAGGAGTACGGTGTCCCTTCTGGAGACAGCGCCCAAGACTGGCATGAATG	581
DB	492	TTGGCCAGGAGTACGGTGTCCCTTCTGGAGACAGCGCCCAAGACTGGCATGAATG	551
QY	582	GAGTAGCCTTCTTGGCCATCGCCAAAGAACTGAATATCCGGCGCGGATCAGCGGAT	641
DB	552	GAGTAGCCTTCTTGGCCATCGCCAAAGAACTGAATATCCGGCGCGGATCAGCGGAT	611
QY	642	GAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAAAGCGCTCCAGCTCTGC	701
DB	612	GAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAAAGCGCTCCAGCTCTGC	671

QY	702	TCCTTCATGTGAATCCAGGGGCGACAGAGAGGCTCTTGAGGACACACAGGATGACGCT	761
DB	672	TCCTTCATGTGAATCCAGGGGCGACAGAGAGGCTCTTGAGGACACACAGGATGACGCT	731
QY	762	TCCTTCATGTGAATCCAGGGGCGACAGAGAGGCTCTTGAGGACACACAGGATGACGCT	821
DB	732	TCCTTCATGTGAATCCAGGGGCGACAGAGAGGCTCTTGAGGACACACAGGATGACGCT	791
QY	822	CTCAGTGTACAGCGCTTCTTCTAGCAGGAGCTATCTTCAACTCTCTTCTTCTTCTTCT	881
DB	792	CTCAGTGTACAGCGCTTCTTCTAGCAGGAGCTATCTTCAACTCTCTTCTTCTTCTTCT	851
QY	882	GGTCTCCCGCATPCCACAGGAGGGTAAACACTTACCTTATTTATTTATTTATTTATTT	941
DB	852	GGTCTCCCGCATPCCACAGGAGGGTAAACACTTACCTTATTTATTTATTTATTTATTT	911
QY	942	TTAATACCAAAAAGGCGCTGATCCCAAAAACCGAGGCTGGGAGCTAGTGGCCCTT	1001
DB	912	TTAATACCAAAAAGGCGCTGATCCCAAAAACCGAGGCTGGGAGCTAGTGGCCCTT	971
QY	1002	TTGCTTTTGTAGGACTTTGGGGCGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1061
DB	972	TTGCTTTTGTAGGACTTTGGGGCGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1031
QY	1062	TCAGCTCAGCCCGAGGGACACAGATGACATTTGGGGGTGGGGGAGGTAAATGACTCCA	1121
DB	1032	TCAGCTCAGCCCGAGGGACACAGATGACATTTGGGGGTGGGGGAGGTAAATGACTCCA	1091
QY	1122	TGCACCCCTCAGTTTCTAGCTGGACAGAGCTCAGGTGACCCAGCCCTTCTTCTTCTTCT	1181
DB	1092	TGCACCCCTCAGTTTCTAGCTGGACAGAGCTCAGGTGACCCAGCCCTTCTTCTTCTTCT	1151
QY	1182	TCTCCAGTGTATCTTCTGCCCCATCTCCCAATTAAGTGGGCCCTTGTGCTGTAGGAA	1241
DB	1152	TCTCCAGTGTATCTTCTGCCCCATCTCCCAATTAAGTGGGCCCTTGTGCTGTAGGAA	1211
QY	1242	GACCAAAAGCTCAGGGAAGATAGAGATATGGAGATGGAGGGGAGGACAAAGGGCAGA	1301
DB	1212	GACCAAAAGCTCAGGGAAGATAGAGATATGGAGATGGAGGGGAGGACAAAGGGCAGA	1271
QY	1302	GAGTAGGCTGTAGCTGGCTATCTTCTGCGCTTACTTACA - CCCCCCTGGAGGATGCCCT	1360
DB	1272	GAGTAGGCTGTAGCTGGCTATCTTCTGCGCTTACTTACA - CCCCCCTGGAGGATGCCCT	1331
QY	1361	TTTCTCCAGCACAAAGACATTTGGGGCACCTGGAATATTTGGTTCAGGCTCCTGTCT	1420
DB	1332	TTTCTCCAGCACAAAGACATTTGGGGCACCTGGAATATTTGGTTCAGGCTCCTGTCT	1391
QY	1421	CTGGACTCAGATCTTCTTCTGCGCCAGGCTGAGGTGAGTAGTGAATCTCCGCTCAC	1480
DB	1392	CTGGACTCAGATCTTCTTCTGCGCCAGGCTGAGGTGAGTAGTGAATCTCCGCTCAC	1451
QY	1481	CTGGCTGTGCACCTTAAACCTCAGGTGAGTAGTGAATCTCCGCTCAC	1540
DB	1452	CTGGCTGTGCACCTTAAACCTCAGGTGAGTAGTGAATCTCCGCTCAC	1511
QY	1541	TTGAATGTGAGTCTTCTTCTGCGCCAGGCTGAGGTGAGTAGTGAATCTCCGCTCAC	1600
DB	1512	TTGAATGTGAGTCTTCTTCTGCGCCAGGCTGAGGTGAGTAGTGAATCTCCGCTCAC	1571
QY	1601	TACAACCTTCACTTCTTCTGCGCCAGGCTGAGGTGAGTAGTGAATCTCCGCTCAC	1660
DB	1572	TACAACCTTCACTTCTTCTGCGCCAGGCTGAGGTGAGTAGTGAATCTCCGCTCAC	1631
QY	1661	GACTATAGTGTGTACCATCACACCTTGGCTTAAATTTTGTATTTTGTAGACACAGGTT	1720
DB	1632	GACTATAGTGTGTGTACCATCACACCTTGGCTTAAATTTTGTATTTTGTAGACACAGGTT	1691
QY	1721	TGCGCATTTGCGCCAGGCTTCTTGAATCTTCTGAGCTCAAGCAACTGCCGCGCTCGGC	1780
DB	1692	TGCGCATTTGCGCCAGGCTTCTTGAATCTTCTGAGCTCAAGCAACTGCCGCGCTCGGC	1751
QY	1781	CTCCCAAGTACTGGGATTACAGCAGAGGACCATGCCCCAGGCTAGATGTGTCTTATC	1840

|||||
Db 1752 CTCCAAAGTACTGGATATACAGCGAGAGGCACCATGCCAGGCTAGATGTCATTATC 1811
QY 1841 CCAATCTTTGGCAGGCTCAGCTCCACAGGCGATTCTTCAAGCAGCTGAAGTGTTTA 1900
Db 1812 CCAATCTTTGGCAGGCTCAGCTCCACAGGCGATTCTTCAAGCAGCTGAAGTGTTTA 1871
QY 1901 GCCTCTCTGGTTAAGAGCCAGATAGAGAAATCCCTTTCTTAGTGTGGAAATGTGTTG 1960
Db 1872 GCCTCTCTGGTTAAGAGCCAGATAGAGAAATCCCTTTCTTAGTGTGGAAATGTGTTG 1931
QY 1961 TG-AAAAAAGAGAAATCCCTGGCTCTGGAGCTGGTGGGAGACAGATTAAGCAAAACC 2019
Db 1932 TGAATAAAGAGAAATCCCTGGCTCTGGAGCTGGTGGGAGACAGATTAAGCAAAACC 1991
QY 2020 TCCCTCTGACATGATCCCTTTGACCCCAAGCTCTGCTCTCTCCCTGACCACCCATGCCCT 2079
Db 1992 TCCCTCTGACATGATCCCTTTGACCCCAAGCTCTGCTCTCTCCCTGACCACCCATGCCCT 2051
QY 2080 TTCTCTTAACCTTCAACAGATACACAGGCGCTAAACTGCTTTACCTCCCTCTACTGA 2139
Db 2052 TTCTCTTAACCTTCAACAGATACACAGGCGCTAAACTGCTTTACCTCCCTCTACTGA 2111
QY 2140 GTCAGGTAGTGGTGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAA 2199
Db 2112 GTCAGGTAGTGGTGGAGGTCACCCATTTCCGAGTTAAACCAATGCAATATGAGTAA 2171
QY 2200 ACAAGTCACTGTGGTATGCTGGGTAGAGAGAGGGTAGCAAGTTTCATGTCTCTCTCT 2259
Db 2172 ACAAGTCACTGTGGTATGCTGGGTAGAGAGAGGGTAGCAAGTTTCATGTCTCTCTCT 2231
QY 2260 TGTCTACATATCCCAAGCTCTGATCCCTGCCATGGGAAGTGGAGAGAAACATGAGG 2319
Db 2232 TGTCTACATATCTCCCAAGCTCCGATCCCTGCCATGGGAAGTGGAGAGAAACATGAGG 2291
QY 2320 TCATGACCTGCAGGCATCTTACTGAGCTCTCCCGCTGGAGGGGAGAGGGGAGGA 2379
Db 2292 TCAATGACCTGCAGGCATCTTACTGAGCTCTCCCGCTGGAGGGGAGAGGGGAGGA 2351
QY 2380 AGAAGTATGCGCTGCACATTTCTGAGGCTACTGCAATTTGCTTTCAAGGCAGAAATCTTG 2439
Db 2352 AGAAGTATGCGCTGCACATTTCTGAGGCTACTGCAATTTGCTTTCAAGGCAGAAATCTTG 2411
QY 2440 TCTGAGCTACAGCGCTCCAGTTTGGCCGATAGAGAGTCTCCGTGGCCCTCCCTCA 2499
Db 2412 TCTGAGCTACAGCGCTCCAGTTTGGCCGATAGAGAGTCTCCGTGGCCCTCCCTCA 2471
QY 2500 GGCAGAGCAGGAGGAGGCTGACATTCGAGTCTCTTCTGGGCCCAAGGAGGTTGCGAG 2559
Db 2472 GGCAGAGCAGGAGGAGGCTGACATTCGAGTCTCTTCTGGGCCCAAGGAGGTTGCGAG 2531
QY 2560 GAGATCCCAATCCCATAGACAGCTCTGGGCTCTTTGATTTTTCAGAAATTAACCT 2619
Db 2532 GAGATCCCAATCCCATAGACAGCTCTGGGCTCTTTGATTTTTCAGAAATTAACCT 2591
QY 2620 GCAGTATTTTGGAAACCAAAAAA 2651
Db 2592 GCAGTATTTTGGAAACCAAAAAA 2623

RESULT 2
AAI60852

ID AAI60852 standard; cDNA; 2323 BP.

XX AC

XX AAI60852;

XX DT

XX 22-OCT-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 4841.

XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

XX KW peripheral nervous system; neuropathy; central nervous system; CNS;

XX KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

XX KW

XX

XX

XX

amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
chemokineti; thrombolytic; drug screening; arthritis; inflammation;
leukaemia; ss.
Homo sapiens.
WO200153312-A1.
XX 26-JUL-2001.
XX 26-DEC-2000; 2000WO-US34263.
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR P-PSDB; AAM41696.
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX Claim 1; SEQ ID NO 4841; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX SQ Sequence 2323 BP; 542 A; 646 C; 613 G; 521 T; 1 other;

Query Match 83.8%; Score 2240; DB 22; Length 2323;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2298; Conservative 0; Mismatches 16; Indels 9; Gaps 4;

QY 15 GGCACCTGCTCAGCTCTCGTCCAGGGACATGACGGGCACGCGCGCTGGCCACCCCG 74
Db 1 GGAATTCCTCAGCTCTCGTCCAGGGACATGACGGGCACGCGCGCTGGCCACCCCG 60
QY 75 GATGGCAGGCGCCCGAGGCGCTCCCGCGCTGCGAGTCCGAGCTACGACCTCAGGGCAAG 134
Db 61 GATGGCAGGCGCCCGAGGCGCTCCCGCGCTGCGAGTCCGAGCTACGACCTCAGGGCAAG 120
QY 135 GTGATGCTTCTGGGAGACACAGCGCTCGGCAAAACATGTTTCTGTATCCAAATCAAGAC 194
Db 121 GTGATGCTTCTGGGAGACACAGCGCTCGGCAAAACATGTTTCTGTATCCAAATCAAGAC 180
QY 195 GGGGCTTCTGTCGGAACCTTCATAGCCACGCTCGGCATAGACTTCAGGAACAAGGTG 254
Db 181 GGGGCTTCTGTCGGAACCTTCATAGCCACGCTCGGCATAGACTTCAGGAACAAGGTG 240

QY 255 GTGACTGTGGATGGCGTGGAGTGAAGTGCAGATCTGGGACACGCTGGCAGGAACGG 314
DB 241 GTGACTGTGGATGGCGTGGAGTGAAGTGCAGATCTGGGACACGCTGGGAGGAACGG 300
QY 315 TTCCGAAGCGTCACCCATGCTTATTTACAGAGATGCTCAGGCGTTCGTCGTGTATGAC 374
DB 301 TTCCGAAGCGTCACCCATGCTTATTTACAGAGATGCTCAGGCGTTCGTCGTGTATGAC 360
QY 375 ATCACAACAATCTCTTTTCACAACATCAGGGCTGGCTCAGTCAGATTCATGAGTAT 434
DB 361 ATCACAACAATCTCTTTTCACAACATCAGGGCTGGCTCAGTCAGATTCATGAGTAT 420
QY 435 GCGCAGAGGCGTGGTGATCATGCTGTAGTGAACAAGCGGATATGAGCAGCGAAAGA 494
DB 421 GCGCAGAGGCGTGGTGATCATGCTGTAGTGAACAAGCGGATATGAGCAGCGAAAGA 480
QY 495 GTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTAGCGGTTCCTTCCTGGAG 554
DB 481 GTGATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTAGCGGTTCCTTCCTGGAG 540
QY 555 ACCAGCGCAAGACTGGCATGAATGTGGAGTTAGGCTTTCTGGCCATCGCCAAGAACTG 614
DB 541 ACCAGCGCAAGACTGGCATGAATGTGGAGTTAGGCTTTCTGGCCATCGCCAAGAACTG 600
QY 615 AAATACCGGGCGGCATCAGGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAG 674
DB 601 AAATACCGGGCGGCATCAGGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAG 660
QY 675 TCCAGAGAGAGCGTCCAGCTGCTGCTTCAATGTGAATCCAGGGGGCAGAGAGAG 734
DB 661 TCCAGAGAGAGCGTCCAGCTGCTGCTTCAATGTGAATCCAGGGGGCAGAGAGAG 720
QY 735 GCTCTGAGGGCACACAGATGCAGCTTCCGCTCCAGGCTGCTTATCCAGAGGC 794
DB 721 GCTCTGAGGGCACACAGATGCAGCTTCCGCTCCAGGCTGCTTATCCAGAGGC 780
QY 795 TGAGCAATGGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTA 854
DB 781 TGAGCAATGGGGAGAAAGATGGAGGACTCACTGCACAGCCGCTTCCTAGCAGGGAGCTA 840
QY 855 TACTCCAACCTCTACTCTGAGTTCTGCGGTCTCCCGCATCCACAGGAGGGTAAACAC 914
DB 841 TACTCCAACCTCTACTCTGAGTTCTGCGGTCTCCCGCATCCACAGGAGGGTAAACAC 900
QY 915 TTAGCTTTTATTTAATAGTACATAATTTAATACCAAAAAGCGCCTGGATCCCAAAA 974
DB 901 TTAGCTTTTATTTAATAGTACATAATTTAATACCAAAAAGCGCCTGGATCCCAAAA 960
QY 975 AACCGAGGCTGGAGCTAGTGGCCCTTTTGTCTTTCTAGGACTTGGGGGCGCGCCCTCCC 1034
DB 961 AACCGAGGCTGGAGCTAGTGGCCCTTTTGTCTTTCTAGGACTTGGGGGCGCGCCCTCCC 1020
QY 1035 TCCTAAGCATAAACAAGTGTGTGCTCCAGCTCAGCCCGAGGGGACACAGATGCATTT 1094
DB 1021 TCCTAAGCATAAACAAGTGTGTGCTCCAGCTCAGCCCGAGGGGACACAGATGCATTT 1080
QY 1095 TGGGGGTGAGGCGAGGTATGACTCCATCGCACCTCAGTTTCAGCTGGACAGGCTCAG 1154
DB 1081 TGGGGGTGAGGCGAGGTATGACTCCATCGCACCTCAGTTTCAGCTGGACAGGCTCAG 1140
QY 1155 GTGACCCCGAGCCTTCACTGTCTCCGCTCTCCAGGAGCTTATCTTGGCCCGCATCTCCCAA 1214
DB 1141 GTGACCCCGAGCCTTCACTGTCTCCGCTCTCCAGGAGCTTATCTTGGCCCGCATCTCCCAA 1200
QY 1215 ATAAGTGGCCCTTGTGCTGTGAGGAAGACCAAGCCCTCAGGGAAGATATGGA 1274
DB 1201 ATAAGTGGCCCTTGTGCTGTGAGGAAGACCAAGCCCTCAGGGAAGATATGGA 1260
QY 1275 GATGGAGGGGAGCACAGGGGACAGAGTAGGCTCTAGCTGGCTATCTCTGGCCCTTAC 1334
DB 1261 GATGGAGGGGAGCACAGGGGACAGAGTAGGCTCTAGCTGGCTATCTCTGGCCCTTAC 1320

QY 1335 TAACACCCCGCTGGAGGATGCCCGCTTTTCTCCAGACACAAAGACATTTGGGGCACCTGG 1394
DB 1321 TAACACCCCGCTGGAGGATGCCCGCTTTTCTCCAGACACAAAGACATTTGGGGCACCTGG 1380
QY 1395 AAATATTGGTTCCAGGCTCCCTGTTCTCTGGACTTCAGATCCTGGGGAGCCCTCCCCC 1454
DB 1381 AAATATTGGTTCCAGGCTCCCTGTTCTCTGGACTTCAGATCCTGGGGAGCCCTCCCCC 1440
QY 1455 CCTGAATCCCTGGCTTAGCTACCTTCCCTGTGCACCTAAAAACCTCAGTTCAGAACT 1514
DB 1441 CCTGAATCCCTGGCTTAGCTACCTTCCCTGTGCACCTAAAAACCTCAGTTCAGAACT 1500
QY 1515 AGGAAAGAGTTTCTTTTATTTTAAATGAGTCTCGTTCTGTGCCCGCAGGCTGA 1574
DB 1501 AGGAAAGAGTTTCTTTTATTTTAAATGAGTCTCGTTCTGTGCCCGCAGGCTGA 1560
QY 1575 GGTGAGTAGTGCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCT 1634
DB 1561 GGTGAGTAGTGCAATCTCCGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCCT 1620
QY 1635 CCACCTCAGCGCGCGAAGTAGCTGGGACTATAGTGTGTACCATCACACCTGGGCTAAAT 1694
DB 1621 CCACCTCAGCGCGCGAAGTAGCTGGGACTATAGTGTGTACCATCACACCTGGGCTAAAT 1680
QY 1695 TTTGTATTTTGTAGACACAGGGTTTCGCCATGTTGCCAGGCTGGTCTTGAATTCCTG 1754
DB 1681 TTTGTATTTTGTAGACACAGGGTTTCGCCATGTTGCCAGGCTGGTCTTGAATTCCTG 1740
QY 1755 AGCTCAAGCAACCTCGCGGCTCGGCTCCCAAGTAGTGTGGATTACAGCGAAGAGGCAC 1814
DB 1741 AGCTCAAGCAACCTCGCGGCTCGGCTCCCAAGTAGTGTGGATTACAGCGAAGAGGCAC 1800
QY 1815 CATGCCAGGCTAGATGTCTTATCCCAATCTTTGGCAGGCATGCAGCTCCACAGGG 1874
DB 1801 CATGCCAGGCTAGATGTCTTATCCCAATCTTTGGCAGGCATGCAGCTCCACAGGG 1860
QY 1875 APTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCTCGGGTTTAAGAGCCAGATAAGGAGAAAT 1934
DB 1861 APTTCTTCAAGCAGCTGAAGTGTTTAGCCCTCTCGGGTTTAAGAGCCAGATAAGGAGAAAT 1920
QY 1935 CCCTTCTCTAGTGTGGATGTGTTGNAAAAAAGAGAAATCCCTGCTGCTGGAGCT 1994
DB 1921 CCCTTCTCTAGTGTGGATGTGTTGNAAAAAAGAGAAATCCCTGCTGCTGGAGCT 1980
QY 1995 GGTGGAGACAAAGATTAAAGCAACCTCCCTGACATGTATCCCTTTGACCCCAAGCTCTG 2054
DB 1981 GGTGGAGACAAAGATTAAAGCAACCTCCCTGACATGTATCCCTTTGACCCCAAGCTCTG 2040
QY 2055 CCTCTCCTGTGACCACCCATGCCCTTTCCTTTAACTTCTCAAAACAGATACCAAGGCGCTAA 2114
DB 2041 CCTCTCCTGTGACCACCCATGCCCTTTCCTTTAACTTCTCAAAACAGATACCAAGGCGCTAA 2100
QY 2115 ACTGCTTTTACCTCCCTCTACTGAGTCAAGTTAGTGTGGGGAGG---TCACCCATTTTC 2171
DB 2101 ACTGCTTTTACCTCCCTCTACTGAGTCAAGTTAGTGTGGGGAGGTTCAACCCATTTTC 2160
QY 2172 CGAGTT---AAACCAATGCAATA-TGAGTAAAAAAGT---CATGTGGGTATGTTCTGGGG 2225
DB 2161 CGAGTTTAAACCAATGCNATATTGAGTAAAAAAGATTCAATGTGGGGTATGTTCTGGGG 2220
QY 2226 TAGAGAGGGGTAGCAAGTTTCATGTCTCTCTTTGGTGCACATATCTCCCAAGCTCTGA 2285
DB 2221 TAGAGAGGGGTAGCAAGTTTCATGTCTCTCTTTGGTGCACATATCTCCCAAGCTCTGA 2280
QY 2286 TCCCTGCATGGGAAGTGGACAGGAACATGAGGTCATGACCT 2328
DB 2281 TCCCTNCCATGGGAAGTGGACAGGAACATGAGGTCATGACCT 2323

RESULT 3
AAS27458
ID AAS27458 standard; cDNA; 1316 BP.
XX

AC AAS27458;
XX
XX DT 07-NOV-2001 (first entry)
XX
DE cDNA encoding novel signal transduction pathway protein, Seq ID 493.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
XX acquired immune deficiency syndrome.
OS Homo sapiens.
PN
XX WO200154733-A1.
XX
XX PD 02-AUG-2001.
XX
XX PF 17-JAN-2001; 2001WO-US01312.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184564.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 03-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 14-SEP-2000; 2000US-0233066.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.

RESULT 5

AAH75184

ID AAH75184 standard; DNA; 1109 BP.

XX

AC AAH75184;

XX

DT 13-NOV-2001 (first entry)

XX

DE Nucleotide sequence of human 32712 G-protein.

XX

Human; G-protein; 32705; 23224; 27423; 32700; 32712; lung disorder; congenital anomaly; pulmonary congestion; oedema; haemorrhage; adult respiratory distress syndrome; Goodpasture's syndrome; chronic obstructive pulmonary disease; asthma; pulmonary hypertension; liver disorder; hepatic injury; jaundice; cholestasis; viral hepatitis; cirrhosis; Wilson's disease; autoimmune hepatitis; hepatic failure; brain disorder; hypoxia; cerebral ischemia; intracranial haemorrhage; acute meningitis; Parkinson's disease; Alzheimer's disease; glioma; chronic bacterial meningoencephalitis; multiple sclerosis; amytropic lateral sclerosis; stroke; Huntington's disease; ss.

OS Homo sapiens.

XX

PH Key Location/Qualifiers

FT CDS

FT 124..699

FT /*tag= a

FT /product= "G-protein"

XX

PN WO200164887-A2.

XX

XX 07-SEP-2001.

XX

XX 27-FEB-2001; 2001WO-US06292.

XX

XX 29-FEB-2000; 2000US-0185606.

XX

XX (MILL-) MILLENNIUM PHARM INC.

XX

XX PA

XX PI

XX MEYERS RA;

XX

DR WPI; 2001-550182/61.

XX

XX P-PSDB; AAG67156.

XX

XX Novel human small G-protein polypeptides and polynucleotides for

PT treating lung disorders, liver disorders and brain disorders -

PT

XX Claim 2; Fig 26; 151pp; English.

XX

The present sequence encodes a human G-protein. The specification describes 32705, 23224, 27423, 32700 or 32712 small G-proteins. The G-protein polypeptides and polynucleotides are useful as a target for diagnosis and treatment of G-protein mediated or related disorders, and for identifying agonists and antagonists for diagnosis and treatment. They are useful for treating disorders of lung (e.g. congenital anomalies, pulmonary congestion, oedema, adult respiratory distress syndrome, haemorrhage, chronic obstructive pulmonary disease, asthma, Goodpasture's syndrome and pulmonary hypertension), liver (e.g. hepatic injury, jaundice, cholestasis, viral hepatitis, cirrhosis, Wilson's disease, autoimmune hepatitis and hepatic failure), and brain (e.g. hypoxia, cerebral ischemia, intracranial haemorrhage, acute meningitis, Parkinson's disease, Alzheimer's disease, gliomas, chronic bacterial meningoencephalitis, multiple sclerosis, amytropic lateral sclerosis, stroke and Huntington's disease).

XX Sequence 1109 BP; 259 A; 309 C; 321 G; 220 T; 0 other;

XX

Query Match

Best Local Similarity

Matches 1082; Conservative

0; Mismatches

5; Indels

0; Gaps

0;

XX

QY 37 GGCATGACGGGACGCCAGGCGCGTGGCCACCGGGATGGCGAGGCCCCGAGCGCT 96

Db	23	CGCTCCGGACGGGACGCCAGGCGCGTGGCCACCCGGATGGCGAGGCCCGGAGCGCT	82
QY	97	CCCCGGCTGCGAGTCCGAGCTAGGACCTCAGGGGCAAGGTGATGCTTCTGGGAGACACAG	156
Db	83	CCCCGGCTGCGAGTCCGAGCTAGGACCTCAGGGGCAAGGTGATGCTTCTGGGAGACACAG	142
QY	157	GGGTGGCAAAACATGTTTCTCTGATCCAAATCAAAGACGGGGCTTCTCTGTCGGGAACCT	216
Db	143	GGGTGGCAAAACATGTTTCTCTGATCCAAATCAAAGACGGGGCTTCTCTGTCGGGAACCT	202
QY	217	TCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTACTGTGGATGGCGTGAGAG	276
Db	203	TCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGGTACTGTGGATGGCGTGAGAG	262
QY	277	TGAAGCTGAGATCTGGGACACCGCTGGCGAGGAACGGTTCGGAAGCGTCACCCATGCTT	336
Db	263	TGAAGCTGAGATCTGGGACACCGCTGGCGAGGAACGGTTCGGAAGCGTCACCCATGCTT	322
QY	337	ATTACAGAGATGCTCAGGCGCTTGTCTGTGATGACATCACCACAATCTTCTTTTCG	396
Db	323	ATTACAGAGATGCTCAGGCGCTTGTCTGTGATGACATCACCACAATCTTCTTTTCG	382
QY	397	ACAACATCAGGCGCTGCTCAGTATGATGATATCCAGAGAGGAGCGTGGTGATCA	456
Db	383	ACAACATCAGGCGCTGCTCAGTATGATGATATCCAGAGAGGAGCGTGGTGATCA	442
QY	457	TGCTGCTAGGCAACAAGCGGGATATGAGCAGCGAAGAGTATCCGTCGGAAGAGCGGAG	516
Db	443	TGCTGCTAGGCAACAAGCGGGATATGAGCAGCGAAGAGTATCCGTCGGAAGAGCGGAG	502
QY	517	AGACCTTGGCCAGGAGTACGGTGTCTCTCTGGAGACCGAGCGCAAGACTGGCATGA	576
Db	503	AGACCTTGGCCAGGAGTACGGTGTCTCTCTGGAGACCGAGCGCAAGACTGGCATGA	562
QY	577	ATGTGAGTATGCTTCTTGGCCATCGCAAGGAACGAAATACCGGGCGGCGCATCAGG	636
Db	563	ATGTGAGTATGCTTCTTGGCCATCGCAAGGAACGAAATACCGGGCGGCGCATCAGG	622
QY	637	CGGATCAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAGAGCGCTCCAGCT	696
Db	623	CGGATCAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAGAGCGCTCCAGCT	682
QY	697	GCTGCTCTTCATGTGAATCCAGGGGGGAGAGAGGAGGCTCTGGAGGCACACAGGATGC	756
Db	683	GCTGCTCTTCATGTGAATCCAGGGGGGAGAGAGGAGGCTCTGGAGGCACACAGGATGC	742
QY	757	AGCCTTCCCTCCAGGCGCTGGCTTATTCGAAGAGGCTGAGCCAATGGGGAGAAAGATG	816
Db	743	AGCCTTCCCTCCAGGCGCTGGCTTATTCGAAGAGGCTGAGCCAATGGGGAGAAAGATG	802
QY	817	GAGGACTCACTGCACAGCGCTTCTTAGCAGGAGCTATCTCAACTCCTACTTGGATT	876
Db	803	GAGGACTCACTGCACAGCGCTTCTTAGCAGGAGCTATCTCAACTCCTACTTGGATT	862
QY	877	CCTGCGGTCTCCCGCATCCACAGGAGGGTAAACACTTTAGCTTTTATTAATAGTAC	936
Db	863	CCTGCGGTCTCCCGCATCCACAGGAGGGTAAACACTTTAGCTTTTATTAATAGTAC	922
QY	937	ATAAATTAATACAAAAAGGCGCTTGGATCCCCAAAAACCGAGGCTGGGAGCTAGTGG	996
Db	923	ATAAATTAATACAAAAAGGCGCTTGGATCCCCAAAAACCGAGGCTGGGAGCTAGTGG	982
QY	997	CCCTTTTGTCTTAGGACTTGGGGGGCGGCGCTCCCTCTAAGCATATCAAGGTGGT	1056
Db	983	CCCTTTTGTCTTAGGACTTGGGGGGCGGCGCTCCCTCTAAGCATATCAAGGTGGT	1042
QY	1057	GTGTCTCAGCTCAGCCCGAGGAGACAGATGCTCTTTGGGGTGGAGGAGGTATGA	1116
Db	1043	GTGTCTCAGCTCAGCCCGAGGAGACAGATGCTCTTTGGGGTGGAGGAGGTATGA	1102
QY	1117	CTCCATC 1123	

OS Homo sapiens.
XX WO200032630-A2.
XX 08-JUN-2000.
XX 06-DEC-1999; 99WO-US28773.
XX 04-DEC-1998; 98US-01111006.
XX (SEAR) SEARLE & CO G D.
XX Dotson SB, Ma X;
XX WPI; 2000-465041/40.
XX Novel nucleic acids derived from activated eosinophil cells useful for
XX treating allergic diseases such as asthma comprises a specific
XX nucleotide sequence -
XX
XX Claim 1; Page 95; 125pp; English.
XX
XX The present invention relates to a number of nucleotide sequences
XX which encode proteins involved in the activation of eosinophils
XX (AAA52601-A52679). Eosinophils are involved in immune reactions, and
XX these genes and their proteins provide possible targets for new drugs to
XX combat diseases such as asthma, allergic rhinitis, atopic dermatitis,
XX anaphylaxis, allergic bronchopulmonary aspergillosis, eczema, psoriasis,
XX emphysema, leukaemia, lymphomas, ovarian cancer, pneumonia, immune
XX disorders, some connective tissue disorders, and inflammatory conditions
XX including septic shock, arthritis, nephritis, inflammatory bowel disease
XX and Crohn's disease.
XX
XX Sequence 797 BP; 192 A; 201 C; 209 G; 195 T; 0 other;
XX
XX Query Match 29.7%; Score 795.4; DB 21; Length 797;
XX Best Local Similarity 99.1%; Pred. No. 3.9e-143;
XX Matches 796; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1840 CCCATCTCTTGGCAGGCATGCGCTCCACAGCGGATTTCTTCAAGCAGCTGAAGTGTT 1899
DB 1 CCCATCTCTTGGCAGGCATGCGCTCCACAGCGGATTTCTTCAAGCAGCTGAAGTGTT 60
QY 1900 AGCCCTCTCTGGGTAAAGACAGATGAAGGAGAAATCCCTTCTAGGTTTGAATGTGT 1959
DB 61 AGCCCTCTCTGGGTAAAGACAGATGAAGGAGAAATCCCTTCTAGGTTTGAATGTGT 120
QY 1960 GTGAAAAAAGAGAAATCCCTGGCTCTGGAGCTGGTGGGAGACAAGATTAAAGCAAAAC 2019
DB 121 GTGAAAAAAGAGAAATCCCTGGCTCTGGAGCTGGTGGGAGACAAGATTAAAGCAAAAC 180
QY 2020 TCCCTGACATGATATCCCTTTGACCCCAAGCTCTGCTCTCCCTGACCAACCATGCCCT 2079
DB 181 TCCCTGACATGATATCCCTTTGACCCCAAGCTCTGCTCTCCCTGACCAACCATGCCCT 240
QY 2080 TTCTTTAACTTCTCAACAGATACACAGGCGCTTAACTGCTTTACCTCCCTCCCTACTGA 2139
DB 241 TTCTTTAACTTCTCAACAGATACACAGGCGCTTAACTGCTTTACCTCCCTCCCTACTGA 300
QY 2140 GTCAGTTTAGTGGTGGGAGGTACCCATTTCCGAGTTAAACCAATGCAATATGAGTAA 2199
DB 301 GTCAGTTTAGTGGTGGGAGGTACCCATTTCCGAGTTAAACCAATGCAATATGAGTAA 360
QY 2200 ACAAAGTCATGCGGTATGCTGGGTAGAGAGAGGGGTAGCAAGTTATGTCCTCTCT 2259
DB 361 ACAAAGTCATGCGGTATGCTGGGTAGAGAGAGGGGTAGCAAGTTATGTCCTCTCT 420
QY 2260 TGCTACATATCTCCCAAGCTCTGATCCCTGCCATGGGAGTGGACAGCAACATGAGG 2319
DB 421 TGCTACATATCTCCCAAGCTCTGATCCCTGCCATGGGAGTGGACAGCAACATGAGG 480
QY 2320 TCATGACCTGCAGGCATCTTACTGTCAGCTCTGCCGCTTGGAGGGGAGAGGGGAGGA 2379
|||||

DB 481 TCATGACCTGCAGGCATCTTTACTGCGAGCTCTGCCGCTGCGAGGGGAGAGGGGAGGA 540
QY 2380 AGAAGTATCGCTGCACATTTCTGAGGCTACTGCAATTTGCTTTCAAGGCAGAAATCTTGC 2439
DB 541 AGAAGTATCGCTGCACATTTCTGAGGCTACTGCAATTTGCTTTCAAGGCAGAAATCTTGC 600
QY 2440 TCTGAGCAGTCAAGCGCTCCAGTTTGGGCCCGGATAGGAAGTTCTCCGTGGCCTCCCTCA 2499
DB 601 TCTGAGCAGTCAAGCGCTCCAGTTTGGGCCCGGATAGGAAGTTCTCCGTGGCCTCCCTCA 660
QY 2500 GGCAGAGCAGGAGGAGGCTGACATTTGCCAGTCTCTCTGGGCCCAAGCAGGTTGCAG 2559
DB 661 GGCAGAGCAGGAGGAGGCTGACATTTGCCAGTCTCTCTGGGCCCAAGCAGGTTGCAG 720
QY 2560 GAGATCCCAATCCATAGACAGCTCTGGGCTCTTGCATTTGAGTTTTCAGAAATTAAC 2619
DB 721 GAGATCCCAATCCATAGACAGCTCTGGGCTCTTGCATTTGAGTTTTCAGAAATTAAC 780
QY 2620 GCAGTATTTTGGAAAGC 2636
DB 781 GCAGTATTTTGGAAAGC 797
RESULT 8
ABA02774
ID ABA02774 standard; DNA; 576 BP.
XX ABA02774;
AC ABA02774;
XX
DT 07-FEB-2002 (first entry)
XX
DE Human degranulation regulator encoding DNA SEQ ID NO 3.
XX
KW Degranulation; mast cell; human; mouse; antiallergic; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..576
FT */*tag= a*
FT */product= "degranulation regulator"*
XX
PN WO200179478-A1.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-JP03268.
XX
PR 19-APR-2000; 2000JP-0118408.
XX
PA (DAIN) DAINIPPON PHARM CO LTD.
XX
PI Yamada T, Ido M;
XX
DR WPI; 2002-041335/05.
XX
DR P-PSDB; AAM52190.
XX
PT Mast cell degranulation controller for treatment of allergies -
XX
PS Claim 11; Page 64-65; 85pp; Japanese.
XX
CC The invention relates to a protein for regulating degranulation of mast
XX cells (degranulation regulators) and the encoding polynucleotides, with
XX antiallergic activity, used in the treatment of allergies associated
XX with degranulation of mast cells.
XX
SQ Sequence 576 BP; 138 A; 147 C; 168 G; 123 T; 0 other;
Query Match 21.2%; Score 568; DB 24; Length 576;
Best Local Similarity 99.1%; Pred. No. 1.2e-99;
Matches 571; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 138 ATGCTTCTGGGAGACAGCGCTCGCAAAACATGTTTCTCTGATCCAATTCAAAGACGG 197

Db 1 ATGCTTCTGGAGCTCGGCGTCGCAAAACATGTTTCTGTATCAAAAGACGGG 60
QY 198 GCCTTCTGTCGGAACTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGTGGTG 257
Db 61 GCCTTCTGTCGGAACTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGTGGTG 120
QY 258 ACTGTGATGGGTGAGAGTGAAGTGCAGATCTGGACACCGCTGGCGAGGAACGGTTC 317
Db 121 ACTGTGATGGGTGAGAGTGAAGTGCAGATCTGGACACCGCTGGCGAGGAACGGTTC 180
QY 318 CGAAGCTCACCATGCTTATTACAGAGATGCTCAGGCGCTTGTCTGTATGACATC 377
Db 181 CGAAGCTCACCATGCTTATTACAGAGATGCTCAGGCGCTTGTCTGTATGACATC 240
QY 378 ACCAACAAATCTCTTTCGACACATCAGGCGCTGCTCACTGAGATTATGATGATGCC 437
Db 241 ACCAACAAATCTCTTTCGACACATCAGGCGCTGCTCACTGAGATTATGATGATGCC 300
QY 438 CAGAGGACGTGTGATGCTGCTAGGCAACAAGCGCGATATGACGCGAAAGATG 497
Db 301 CAGAGGACGTGTGATGCTGCTAGGCAACAAGCGCGATATGACGCGAAAGATG 360
QY 498 ATCGTTCCGAAGCAGGAGACCTTGGCCAGGAGTACGGTGTTCCTTCTGGAGACC 557
Db 361 ATCGTTCCGAAGCAGGAGACCTTGGCCAGGAGTACGGTGTTCCTTCTGGAGACC 420
QY 558 AGCGCCAAAGACTGGCATGATGTGGAGTGTAGCCCTTCTGGCCATCGCAAGGAACGAAA 617
Db 421 AGCGCCAAAGACTGGCATGATGTGGAGTGTAGCCCTTCTGGCCATCGCAAGGAACGAAA 480
QY 618 TACCGGCGCGGCATCAGCGGATGAGCCAGGTTCAGATCCAGACTATGATGAGTCC 677
Db 481 TACCGGCGCGGCATCAGCGGATGAGCCAGGTTCAGATCCAGACTATGATGAGTCC 540
QY 678 CAGAAGAGCGCTCCAGCTGCTCTCTCATGTGA 713
Db 541 CAGAAGAGCGCTCCAGCTGCTCTCTCATGTGA 576

RESULT 9
AA86720
ID AA86720 standard; DNA; 843 BP.
XX AC AA86720;
XX 27-OCT-1999 (first entry)
XX DNA sequence encoding an exocytotic protein designated Exo2.

XX Exocytotic protein; Exo2; exocytosis; Chediak-Higashi Syndrome;
KW Inflammation; ss.

XX Mus sp.
FH Key Location/Qualifiers
FT CDS 124..699
FT FT /*tag= a
FT FT /transl_except= (pos: 136..138, aa: xaa)
FT FT /transl_except= (pos: 202..204, aa: xaa)
FT FT /note= "Xaa is an unspecified amino acid"

XX WO9942586-A2.
XX 26-AUG-1999.
XX 23-FEB-1999; 99WO-US03909.
XX 26-MAY-1998; 98US-0086650.
XX 23-FEB-1998; 98US-0075534.
XX (RIGE-) RIGEL PHARM INC.

PI Anderson D, Fisher J, Huang B, Lorens J, Luo Y;
PI Shen M;
XX
DR WPI; 1999-518605/43.
DR P-PSDB; AAY30132.
XX
PT New exocytotic proteins useful for diagnosis and treatment of
PT exocytosis-mediated conditions and in drug screening
XX
PS Claim 6; Fig 6; 53pp; English.

XX The present sequence encodes an exocytotic protein designated Exo2.
CC The protein has an inhibitory effect on exocytosis (i.e. the fusion of
CC secretory vesicles with the cellular plasma membrane) and can be
CC administered therapeutically to treat or prevent exocytosis-mediated
CC disorders e.g. Chediak-Higashi Syndrome (CHS) or conditions involving
CC undesirable release of compounds via exocytosis e.g. inflammation
CC mediated by the release compounds such as histamine. Exo2 is useful
CC diagnostically and to produce antibodies useful to purify the proteins
CC and therapeutically to reduce or eliminate the biological activity of
CC the protein. Exo2 nucleic acids can be used therapeutically to increase
CC Exo2 activity in cells by known gene therapy techniques. They can also
CC be used to produce probes or primers to isolate Exo2 proteins from other
CC organisms, especially humans. The nucleic acids, host cells and proteins
CC are useful in screening assays to identify binding agents, especially
CC drug screening assays to identify agonists and antagonists useful
CC therapeutically to enhance or reduce Exo2 activity.

XX Sequence 843 BP; 195 A; 218 C; 239 G; 175 T; 16 other;

Query Match 20.6%; Score 550.8; DB 20; Length 843;
Best Local Similarity 84.9%; Pred. No. 2.5e-96;
Matches 635; Conservative 0; Mismatches 110; Indels 3; Gaps 2;

QY 32 GTCACGGGACATCAGCGGCACGCCAGGCGC--CGTTGCCACCCCGGATGCGAGGCCCC 89
Db 16 GNNCGGNGANATGAGCGACACAGAGGCTTGCTACCGCTGGGATGCGAGGCCCT 75

QY 90 GAGCGCTCCCGCCCTGCGAGTCCGAGCTACGACCTCAGCGGCAAGGTGATGCTTCGGGA 149
Db 76 GAGCGCTCCCGCCCTCAGCGGCAAGGTGATGCTTCGGGA 135

QY 150 GACACAGGCGTCGGCAAAACATGTTCCCTGATCCATTCAAAGACGGGCTTCCTGTCC 209
Db 136 GANTCGGGCGTCGGCAAAACCTGTTCCCTGATCCATTCAAAGACGGGCTTCCTGTCC 195

QY 210 GGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGCTGTGATGGC 269
Db 196 GGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTGCTGTGATGGC 255

QY 270 GTGAGAGTGAAGTGCAGATCTGGGACACCGCTGGGAGGAGGAGGAGGAGGAGGAGG 329
Db 256 TCCAGGGGTGAAGCTTCAGATCTGGGACATGCGAGGACAGGAGGAGGAGGAGGAGG 315

QY 330 CATGCTTATTACAGAGATGCTCAGGCGCTTCTGCTGTATGATCATCACAACAAATCT 389
Db 316 CATGCTTATTACAGAGATGCTCAGGCGCTTCTGCTGTATGATCATCACAACAAATCT 375

QY 390 TCTTTTCACAACATCAGGCGCTGCTCAGTACGATTCATGAGTATGCCAGAGGAGGAGG 449
Db 376 TCTTTTCACAACATCAGGCGCTGCTCAGTACGATTCATGAGTATGCCAGAGGAGGAG 435

QY 450 GTGATCATGCTGTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTATGCTTCGGA 509
Db 436 GTGATCATGCTGTAGGCAACAAGGCGGATATGAGCAGCGAAAGAGTATGCTTCGGA 495

QY 510 GACGGAGAGACCTTGGCCAGGAGTACGGTGTTCCTCTTCTGGAGACCGGCAAGACT 569
Db 496 GATGGAGAGACCTTGGCCAGGAGTATGCTTCCTCTTCTGGAGACCGGCAAGACT 555

QY 570 GGCATGAATGTGGAGTGTAGCTTCTTGGCCATCGCCAAAGAACTGAAATACCGGCGGG 629
Db 556 GGCATGAATGTGGAGTGTAGCTTCTTGGCCATCGCCAAAGAACTGAAATACCGGCGGG 615

Qy	630	CATPCAGGCGGATGAGCCCAAGCTTCAGATCCGAGACTATGTAGTCCAGAGAAGCGCC	689
Db	616	AGGCAGCCTGATGAGCCCAAGCTTCAGATCCGAGACTATGTGGAGTCCCGAGAAGCGC	675
Qy	690	TCCAGCTGCTGCTCCCTTCATGTGAATCCACGGGGCGAGAGAGGCTCTCGAGGCAC-A	748
Db	676	TCCAGCTGCTGCTCCCTTTGTGTGACCCCTAGGGGCTAAGAGAGGGCCCAAGACCCCTT	735
Qy	749	CAGGATCCAGCCCTTCCCTCCCGAGGCC	776
Db	736	GGGGATCGAGTACTCCCAACTGCCACACC	763
RESULT 10			
ABA02773	ID	ABA02773 standard; DNA; 576 BP.	
XX	AC	ABA02773;	
XX	DT	07-FEB-2002 (first entry)	
XX	DE	Mouse degranulation regulator encoding DNA SEQ ID NO 2.	
XX	KW	Degranulation; mast cell; human; mouse; antiallergic; ds.	
XX	OS	Mus musculus.	
XX	Key	Location/Qualifiers	
FT	CDS	1..576	
FT		/*tag= a	
FT		/transl_except= (pos:247..249,aa:Lys)	
FT		/product= "degranulation regulator"	
XX	XX		
PN	WO200179478-A1.		
XX	XX		
PD	25-OCT-2001.		
XX	XX		
PF	16-APR-2001; 2001WO-JP03268.		
XX	XX		
PR	19-APR-2000; 2000JP-0118408.		
XX	XX		
PA	(DAIN) DAINIPPON PHARM CO LTD.		
XX	XX		
PI	Yamada T, Ido M;		
XX	XX		
DR	WPI; 2002-041335/05.		
DR	P-PSDB; AAM52189.		
XX	XX		
PT	Mast cell degranulation controller for treatment of allergies -		
XX	XX		
PS	Claim 10; Page 63-64; 85pp; Japanese.		
XX	XX		
CC	The invention relates to a protein for regulating degranulation of mast cells (degranulation regulators) and the encoding polynucleotides, with antiallergic activity, used in the treatment of allergies associated with degranulation of mast cells.		
XX	XX		
SQ	Sequence 576 BP; 139 A; 139 C; 165 G; 133 T; 0 other;		
Query Match 18.0%; Score 480; DB 24; Length 576;			
Best Local Similarity 89.6%; Pred. No. 8.5e-83;			
Matches 516; Conservative 0; Mismatches 60; Indels 0; Gaps			
Qy	138	ATGCTTCTGGAGACACAGCGCTCGGCAAAACATGTTTCTGTATCCAATTCAAAGACGGG	197
Db	1	ATGCTTCTGGAGACTCGGCGCTCGGCAAAACCTGTTTCTGTATCCAATTCAAAGACGGG	60
Qy	198	GCCTTCTGTCGCGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAACAGGTGGTG	257
Db	61	GCCTTCTGTCGCGAACCTTCATAGCCACCGTCGGCATAGACTTCAGGAATAAAGTGGTG	120
Qy	258	ACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGCACACCGCTGGCGAGGAACGGTTC	317

Db	121	ACAGTGGATGGTGCCAGGGTGAAGCTTCAGATCTGGACACTGCAGACAGAGGCCCTTT	181
Qy	318	CGAAGCGTCAACCATGCTTATTACAGAGATGCTCAGGCCCTTGCTTCTGCTATGACATC	377
Db	181	CGCAGTGTGACCCATGCTTATTACCGAGATGCTCAGGCCTTGCTCTGTTGTATGACATC	240
Qy	378	ACCAACAATCTTCTTCTTCGACACATCAGGGCCTGGCTCACTGAGATTATCATGAGTATGCC	437
Db	241	ACCAACCAGTCTCTTTTGTGACAACATCAGGGCCTGGCTCACAGAGATTTCATGAGTATGCC	300
Qy	438	CAGAGGACCTGTGTGATCATGCTGCTAGGCCAACAGCGGATATGAGCAGCGAAACAGTG	497
Db	301	CAGAGAGACCTGTGTGATTAATGCTTCTAGGCCAACAGCGCGATTAAGCAGCGAAAGGGTG	360
Qy	498	ATCCGTTTCCGAAGACGGAGAGACCTTGGCCAGGGAGTACGGTGTTCCTTCTCTGGAGACC	557
Db	361	ATCCGTTCTCAAGATGGAGAGACATGGCCAGGGAATATGGTGTCTCTTCATGGAGACC	420
Qy	558	AGGCCCAAGACTTGGCATGAATGTGGAGTTAGCCTTTCTGCGCCATCGCCAAAGAACTGAAA	617
Db	421	AGTGCCAAGACTTGGCATGAACGTGGAGTTGGCCTTTCTGGCAATTGCCAAAGAACTGAAA	480
Qy	618	TACCGGCCCGGCATCAGGCGGATGAGCCAGCTTCAGATCCGAGCATATGTAGAGTCC	677
Db	481	TACCGTCAAGGAGGACGCTGATGAGCCAGCTTCAGATCCGAGCATATGTGGAGTCC	540
Qy	678	CAGAAGAAGCGCTCCAGCTGCTGCTTTCATGTGA	713
Db	541	CAGAAGAAGCGCTCCAGCTGCTGCTTGTGTGA	576
RESULT 11			
AAS81431			
ID	AAS81431 standard; cDNA; 475 BP.		
AC	AAS81431;		
XX			
DT	13-FEB-2002 (first entry)		
XX			
DE	DNA encoding novel human diagnostic protein #17235.		
XX			
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;		
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	W0200175067-A2.		
XX			
PD	11-OCT-2001.		
XX			
XX	30-MAR-2001; 2001WO-US08631.		
PF			
XX			
PR	31-MAR-2000; 2000US-0540217.		
PR	23-AUG-2000; 2000US-0649167.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Drmanac RT, Liu C, Tang YT;		
XX			
XX	WPI; 2001-639362/73.		
DR	P-PSDB; ABG17244.		
XX			
PT	New isolated polynucleotide and encoded polypeptides, useful in		
PT	diagnostics; forensics, gene mapping, identification of mutations		
PT	responsible for genetic disorders or other traits and to assess		
PT	biodiversity		
XX			
PS	Claim 1; SEQ ID No 17235; 103pp; English.		
XX			
CC	The invention relates to isolated polynucleotide (I) and		
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,		
CC	polymerase chain reaction (PCR) primers, oligomers, and/or chromosomal		

CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 475 BP; 108 A; 136 C; 138 G; 93 T; 0 other;

Query Match 17.5%; Score 468.6; DB 23; Length 475;
Best Local Similarity 99.2%; Pred. No. 1.3e-80;
Matches 471; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 15 GGCACGTCTACCTCTCGTCAGGGACATGACGGGACGCGCGGTGGCCACCGG 74
DB 1 GGAATTCTCTACCTCTCGTCAGGGACATGACGGGACGCGCGGTGGCCACCGG 60
QY 75 GATGGGAGGCCCCCGAGCGCTCCCGCTCGAGTCCGAGCTACGACCTCAGGGCAAG 134
DB 61 GATGGGAGGCCCCCGAGCGCTCCCGCTCGAGTCCGAGCTACGACCTCAGGGCAAG 120
QY 135 GTGATGCTCTGGGAGACACAGCGCTCGGCAAAACATGTTTCTGTATCCAAATCAAGAC 194
DB 121 GTGATGCTCTGGGAGACACAGCGCTCGGCAAAACATGTTTCTGTATCCAAATCAAGAC 180
QY 195 GGGGCTTCTGTCGGGAACCTTATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTG 254
DB 181 GGGGCTTCTGTCGGGAACCTTATAGCCACCGTCGGCATAGACTTCAGGAACAAGGTG 240
QY 255 GTCACTGTGATGCGTGAAGTGAAGTGCAGATCTGGGACACCGCTGGGAGGACCG 314
DB 241 GTCACTGTGATGCGTGAAGTGAAGTGCAGATCTGGGACACCGCTGGGAGGACCG 300
QY 315 TTCGGAAGCGTACCCATGCTTTATACAGAGATGCTCAGGCTTGTCTGCTGTATGAC 374
DB 301 TTCGGAAGCGTACCCATGCTTTATACAGAGATGCTCAGGCTTGTCTGCTGTATGAC 360
QY 375 ATCACCACAAATCTCTTTTCGACACATCAGGGCGCTGGCTCACTGAGATTCATGATAT 434
DB 361 ATCACCACAAATCTCTTTTCGACACATCAGGGCGCTGGCTCACTGAGATTCATGATAT 420
QY 435 GCCCAGAGGACGCTGTGATCATGCTCTAGGCAACAAGCGGATATGACGACCG 489
DB 421 GCCCAGAGGACGCTGTGATCATGCTCTAGGCAACAAGCGGATATGACGACCG 475

RESULT 12

AAV65197
ID AAV65197 standard; DNA; 1340 BP.
XX AC AAV65197;
XX DT 17-DEC-1998 (first entry)
XX DE Human RAB protein, SRAB, coding sequence.
XX KW Human; RAB protein; SRAB; vesicle trafficking disorder; epilepsy; cancer;
KW Huntington's disease; Parkinson's disease; schizophrenia; choroideraemia;
KW viral infection; therapy; autoimmune sialosis; cystic fibrosis;
KW diabetes mellitus; hyperglycaemia; tubulointerstitial nephritis;
KW hypoglycaemia; pancreatic enzyme deficiency; pancreatitis;

KW polycystic renal disease; ss.
XX Homo sapiens.
FH Key Location/Qualifiers
FT CDS 43..615
FT /tag= a
FT /product= "SRAB"
XX W09842839-A1.
XX 01-OCT-1998.
XX 25-MAR-1998; 98WO-US03871.
XX 26-MAR-1997; 97US-0824873.
XX (INCY-) INCYTE PHARM INC.
XX Guegler KJ, Hillman JL;
XX WPI: 1998-542282/46.
XX P-PSDB; AAW80747.
XX New human Rab protein, SRAB - useful for treating disorders
PT associated with SRAB expression, including vesicle trafficking,
PT viral infection, and cancer
XX Claim 5; Fig 1; 31pp; English.
XX This sequence encodes the human RAB protein, SRAB, of the invention. The
CC SRAB protein can be used in compositions useful in the diagnosis,
CC prevention, or treatment of disorders associated with vesicle trafficking
CC (including epilepsy, Huntington's disease, Parkinson's disease and
CC schizophrenia), cancer, or viral infection. Host cells containing the DNA
CC sequence can be used to produce SRAB recombinantly. The viral infections
CC that can be treated using SRAB include those caused by retroviruses
CC particularly HIV and HTLV, hepatitisviruses, particularly hepatitis C,
CC hantaviruses, herpesviruses, and arboviruses. The cancers that can be
CC treated include adenocarcinoma, leukaemia, lymphoma, melanoma and
CC sarcoma, particularly cancers of the endocrine, gastrointestinal and
CC nervous systems and cancers of the adrenal gland, brain, breast, colon,
CC oesophagus, kidney, liver, lung, ovaries, pancreas, pituitary gland,
CC prostate, salivary gland, stomach, thyroid, and uterus. The protein can
CC also be used to treat conditions such as autoimmune sialosis,
CC choroideraemia, cystic fibrosis, diabetes mellitus, hyperglycaemia,
CC hypoglycaemia, tubulointerstitial nephritis, pancreatic enzyme
CC deficiency, pancreatitis, and polycystic renal disease.
XX

XX Sequence 1340 BP; 276 A; 387 C; 403 G; 274 T; 0 other;

Query Match 10.9%; Score 292.6; DB 19; Length 1340;
Best Local Similarity 70.9%; Pred. No. 6.7e-47;
Matches 416; Conservative 0; Mismatches 169; Indels 2; Gaps 2;

QY 116 CTACGACCTCAGCGGCAAGGTGATGCTTCTGGGAGACACAGGCGTCGGCAAAACATGTTT 175
DB 21 CTACGACGTCGCTTCAAGGTATGCTGTGGGGGACGCGGTGGGGAAGACCTGTCT 80
QY 176 CCTGA-TCCAAATCAAGACGGGCGCTTCTGTCCGGACCTTCATAGCCCGTCGGCA 234
DB 81 GCTGGGTGGGATTCAAAGGATGGTGTCTTCTGGCGGGGACCTTCATCTCCACCGT-AGCA 139
QY 235 TAGACTTCAGGAACAAGGTGGTGACTGTGGATGGCTGAGAGTGAAGCTGCAGATCTGGG 294
DB 140 TTGACTTCGGGAACAAGTTCCTGGACGTGGATGGTGTGAAGGTGAAGCTGCAGATCTGGG 199
QY 295 ACACCGCTGGGAGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGG 354
DB 200 ACACAGCTGTCAGGAGCGGTTCCGCGAGTGTACCCTACTGCTACTACCGGGATGCTCATG 259
QY 355 CTTTGGTCTTCTGTGATGACATCACCACAAATCTTCTTCGACACATCAGGGCGCTGGC 414
XX

Db 260 CTCCTGCTGCTACAGATGTCACCAACAAGGCTCTCTTTGACAACTCCAGGCTGGC 319
QY 415 TCACCTGAGATTCATGAGTATGCCAGAGGAGCTGGTGATCATGCTGTAGCAACAAGG 474
Db 320 TGACCGAGATCCACAGTACGCCAGCAGAGCTGGCTCATGCTGTGGGGAACAAGG 379
QY 475 CGGATATGAGCAGCGAAGAGATGATCCGTTCCGAGAGCGGAGACCTTGGCCAGGGAGT 534
Db 380 TGGACTCTGCCATGAGCTGTGGTGAAGAGGGAGGAGCGGGAGAAGCTGGCCAAGGAGT 439
QY 535 ACGGCTTCCTCTCTGAGACCGCGCAAGACTGGCATGAATGGAGTTAGCCTTTC 594
Db 440 ATGGACTGCGCTTCATGGAGACCGCGCAAGAGCGGCTCAACGCTGACTTGGCCTTCA 499
QY 595 TGGCCATGCCCAAGGAAGTGAATACCGCGGCGGCGATCAGCGGATGAGCCAGCTTCC 654
Db 500 CAGCCATAGCAAGAGTGAAGCAGCGCTCCATGAAGGCTCCAGCGAGCGCGCTTCC 559
QY 655 AGNCCGAGACTATGATAGTCCCAAGAGAGCGCTCCAGCTGCTGC 701
Db 560 GGCTGCATGATACGTTAAGAGGAGGCTCGAGGGGCGCTCTCTGCTGC 606

RESULT 13

AAA89717

ID AAA89717 standard; cDNA; 443 BP.

XX

AC AAA89717;

XX

DT 08-JAN-2001 (first entry)

XX

DE Mouse Exo104 nucleotide sequence.

XX

KW Exocytosis pathway protein; Exo protein; antiparkinsonian; antidiabetic;
KW antiallergic; antiasthmatic; nootropic; neuroprotective; anticonvulsant;
KW vulnerable; asthma; inflammation; allergy; Chediak-Higashi syndrome; CHS;
KW Alzheimer's disease; Parkinson's disease; Huntington's disease; diabetes;
KW digestion disorder; wound healing disorder; gene therapy; ss.

XX Mus sp.

OS

XX WO200043419-A2.

PN

XX 27-JUL-2000.

PD

XX 20-JAN-2000; 2000WO-US01431.

XX

PR 20-JAN-1999; 99US-0116534.

PR 26-JAN-1999; 99US-0117274.

PR 26-JAN-1999; 99US-0117308.

PR 26-JAN-1999; 99US-0117312.

PR 01-FEB-1999; 99US-0118177.

PR 01-FEB-1999; 99US-0118178.

PR 01-FEB-1999; 99US-0118179.

PR 09-FEB-1999; 99US-0119286.

PR 11-FEB-1999; 99US-0119998.

PR 11-FEB-1999; 99US-0119759.

XX

PA (RIGE-) RIGEL PHARM INC.

XX

PI Luo Y;

XX

DR WPI; 2000-482908/42.

XX

PT New nucleic acids encoding Exo proteins which are useful in the
PT diagnosis, treatment or prevention of exocytosis-mediated disorders
PT such as asthma, inflammation and allergies -

XX

PS Disclosure; Page 290; 305pp; English.

XX

XX The present sequence encodes a polypeptide which is associated with
CC the exocytosis pathway. cDNA molecules encoding proteins involved in

CC exocytosis have been isolated by yeast one-hybrid and two-hybrid
CC screening. Novel proteins, termed Exo proteins, have been identified that
CC interact with known exocytosis-associated proteins such as GS27, alpha
CC snap, unc18-L, vamp3, snap-23, and the rab family of proteins. the
CC Exo proteins and their agonists and antagonists are useful in the
CC diagnosis, treatment or prevention of exocytosis-mediated disorders
CC such as asthma, inflammation, allergies, Chediak-Higashi Syndrome
CC (CHS), Alzheimer's disease, Parkinson's disease, Huntington's disease,
CC diabetes, digestion disorders and wound healing disorders.
CC The nucleic acids, antagonists or agonists of Exo proteins are useful
CC in gene therapy. The nucleic acids are also useful for generating
CC transgenic or knock-out animals which can be used in the
CC development and screening of therapeutically useful reagents.

XX
SQ Sequence 443 BP; 109 A; 111 C; 131 G; 90 T; 2 other;

Query Match 10.2%; Score 274; DB 21; Length 443;

Best Local Similarity 84.6%; Pred. No. 2.1e-43;

Matches 307; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 414 CTCACCTGAGATTCATGAGTATGCCAGAGGAGCTGGTGATCATGCTGTAGCAACAAG 473
Db 1 CTCACAGATTCATGAGTATGCCAGAGGAGCTGGTGATCATGCTGTAGCAACAAG 60

QY 474 GCGGATATGAGCAGCGAAGAGTGTATCCGTAAGACGGAGAGACCTTGGCCAGGGAG 533
Db 61 GCCGATGTAAGCAGCGAAGAGGTGATCCGTTCTGAAGATGGAGAGACACTGGCCAGGAA 120

QY 534 TAGCGTGTTCCTTCTGGAGACCGCCAAAGACTGGCATGAATGTGGAGTTAGCCTTT 593
Db 121 TATGGTGTTCCTTCTATGGAGACCAAGTGCCTCAAGACTGGCATGAATGTGGAGTTAGCCTTT 180

QY 594 CTGGCCATCGCCCAAGAACTGAATACCGCGCGGCATCAGCGGATCAGCCAGCTTC 653
Db 181 CTGGCAATTCGCCAAGAACTGAATACCGTCGAGGAGGAGCAGCTGATGAGCCAGCTTC 240

QY 654 CAGATCCGAGACTATGTAGAGTCCCAAGAGAGCGCTCCAGCTGTGCTCTTCATGTGA 713
Db 241 CAGATCCGAGACTATGTGGAGTCCCAAGAGAGCGCTCCAGCTGTGCTCTTCATGTGA 300

QY 714 ATCCAGAGGGGAGAGAGGAGCTCTGGAGGAGCACACAGGATCGAGCCTTCCCCCTCCAG 773
Db 301 CCCCTAGGGGTAAGAGGAGGAGGCCANAGACCTTGGGGATGCAGTACTCCAAGTCCAC 360

QY 774 GCC 776

Db 361 ACC 363

RESULT 14

ABL29523

ID ABL29523 standard; DNA; 2001 BP.

XX

AC ABL29523;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 40042.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX

PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 21-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241836.
 PR 01-NOV-2000; 2000US-0244637.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249219.
 PR 17-NOV-2000; 2000US-0249220.
 PR 17-NOV-2000; 2000US-0249221.
 PR 17-NOV-2000; 2000US-0249224.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.

PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPT: 2001-465460/50.
 DR P-PSDB; AAU1123.
 XX
 PT Novel polypeptides useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cancers, immune
 PT disorders and neuronal disorders
 XX
 PS Claim 1; SEQ ID No 75; 880pp; English.
 XX
 CC The invention relates to novel isolated polypeptides (I), and
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
 CC diagnosing, preventing and treating diseases including immune system
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
 CC transplant rejections and graft versus host disease, infectious diseases
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative
 CC disorders, primary haematopoietic disorders, hyperproliferative
 CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
 CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
 CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
 CC Addison's disease), reproductive system disorders, gastrointestinal
 CC disorder (inflammatory disorders), liver disorders (cirrhosis),
 CC as stimulators of B-cell responsiveness to pathogens, activators of
 CC T-cells, to induce higher affinity antibodies, and as a means to induce
 CC tumour proliferation in pathologies e.g. acquired immune deficiency
 CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
 CC pathway protein coding sequences and PCR primers of the invention.
 XX

Query Match 7.8%; Score 209.4; DB 22; Length 964;
 Best Local Similarity 69.3%; Pred. No. 5, 4e-31;
 Matches 285; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
 QY 291 TGGGACACCGTGGCGAGGAAGGTTCCGAAAGCGTCACCCATGCTTATTACAGATGCT 350
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 3 TGGGACACAGCTGGTCCAGGAGCGGTTCCGACGTGTACCCATGCTTACCGGGATGCT 62
 QY 351 CAGGCCTTGTCTCTGTATGACATCAACCAAAATCTTCTTTCACAAATCAGGGCC 410
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 63 CATGCTCTGTCTGTCTGTACGATGTACCAACAAAGGGCTCTTTGACAAATCAGGGCC 122
 QY 411 TGGCTCACTGAGATTCATGATGATCCAGAGGAGCGTGGTGTATCATGCTAGTACGAC 470
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 123 TGGCTGACCGAGATCCAGGATCCAGGATCCGCGGCGCTCATGCTGTGGGGAAC 182
 QY 471 AAGCGGATATGAGCAGGCAAGAGAGTGTATCCCTCCGAGACGAGAGACCTTGCACGG 530
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 183 AAGGTGACTCTGCCCATGAGCGTGTGGTGAAGAGGGAGGACGGGGAGAGAGCTGGCCAAAG 242
QY 531 GAGTACGGTGTTCCTTCCTGGAGACCCAGCGCCAAAGACTGGCATGGAATGTGGAGTTAGCC 590
    ||||| || | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 243 GAGTATGGAAGTGGCCCTTCATGGAGACCCAGCGCCAAAGAGCGGCCCTCAACGTGGACTTGGCC 302
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 591 TTTCGTGGCCATCGCCAAAGGAAGTGAATACCGGGCCGGGCATCAGGCGGATGAGCCCAAGC 650
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 303 TTCACAGCCATAGCAAGGAGTTGAAGCAGCGCTCCATGAAGGTCCTCCAGCGAGCCGCGC 362
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 651 TTCAGATCCGAGACTATGTAGTCCCAAGAGAGCGCTCCAGCTGCTGC 701
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 363 TTCGGCTGCATGATTACGTTAAGAGGAGGTCGAGGGGCCCTCCTGCTGC 413
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Search completed: November 17, 2002, 14:38:28
 Job time : 379 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 17, 2002, 14:30:48 ; Search time 83 Seconds
(without alignments)
11904.524 Million cell updates/sec

Title: US-09-817-199A-1
Perfect score: 2674
Sequence: 1 ttccgctgcggcgccgcact.....aaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 328367 seqs, 18475068 residues
Total number of hits satisfying chosen parameters: 656734

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2674	100.0	2674	10	US-09-817-199A-1
2	2031.8	76.0	13182	10	US-09-817-199A-3
3	1086	40.6	1116	12	US-09-794-257-13
4	873	32.6	875	12	US-10-051-986-10
5	683.4	25.6	24707	10	US-09-740-027-3
6	600.6	22.5	601	10	US-09-817-199A-24
7	576	21.5	576	10	US-09-794-257-15
8	573	21.4	601	10	US-09-817-199A-25
9	443.4	16.6	447	10	US-09-867-550-1811
c 10	200.6	7.5	21833	10	US-09-764-877-2275
c 11	199.8	7.5	1130	10	US-09-764-877-3373
c 12	198.6	7.4	15745	10	US-09-764-877-3236
c 13	198.6	7.4	17216	10	US-09-764-877-3565
c 14	198.6	7.4	17217	10	US-09-764-877-3566
c 15	198.4	7.4	4962	10	US-09-764-869-1902
c 16	198.4	7.4	4962	10	US-09-764-869-1903
c 17	197.6	7.4	7791	10	US-09-764-877-3700
c 18	197.2	7.4	2754	10	US-09-745-763-133
19	197.2	7.4	15515	10	US-09-822-860-3

c 20	196.8	7.4	17335	10	US-09-764-847-1280
c 21	196.8	7.4	19882	10	US-09-764-847-1281
c 22	196.6	7.4	13821	10	US-09-764-877-2595
c 23	196.2	7.3	7960	10	US-09-764-869-2327
c 24	195.8	7.3	203654	10	US-09-820-905-3
c 25	195.4	7.3	30175	10	US-09-738-878-3
c 26	195.4	7.3	30175	12	US-10-163-381-3
c 27	195.2	7.3	17239	10	US-09-764-869-2326
c 28	195.2	7.3	55795	10	US-09-880-107-1543
c 29	195	7.3	13819	10	US-09-764-877-2596
c 30	195	7.3	180557	12	US-10-003-806-6
c 31	195	7.3	180557	12	US-10-003-806-9
c 32	194.6	7.3	32146	10	US-09-764-860-797
c 33	194.6	7.3	32248	10	US-09-764-860-802
c 34	194.4	7.3	32822	10	US-09-964-824A-572
c 35	194.4	7.3	49984	10	US-09-739-457-5
c 36	194.4	7.3	57130	10	US-09-835-081-3
c 37	194.4	7.3	58837	10	US-09-982-091A-5
c 38	194.4	7.3	326014	10	US-09-731-231A-3
c 39	194	7.3	48436	10	US-09-927-602-38
c 40	193.8	7.2	32176	10	US-09-764-869-1953
c 41	193.8	7.2	32178	10	US-09-764-869-1954
c 42	193.6	7.2	10252	10	US-09-764-887-442
c 43	193.6	7.2	24699	10	US-09-764-877-2419
c 44	193.4	7.2	12503	10	US-09-764-870-549
c 45	193.4	7.2	31871	10	US-09-764-847-1403

ALIGNMENTS

RESULT 1
US-09-817-199A-1
; Sequence 1, Application US/09817199A
; Patent No. US20020142380A1
; GENERAL INFORMATION:
; APPLICANT: SHAO, Wei et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO01187
; CURRENT APPLICATION NUMBER: US/09/817,199A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2674
; TYPE: DNA
; ORGANISM: Human
US-09-817-199A-1

Query Match	100.0%	Score 2674;	DB 10;	Length 2674;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2674;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	TTTCGCTGCGGCGCGGCACTGCTCACTCTCGTCCAGGACATGACGGGCACGCCAGGCG	60	
Db	1	TTTCGCTGCGGCGCGGCACTGCTCACTCTCGTCCAGGACATGACGGGCACGCCAGGCG	60	
Qy	61	CGTTTGCCACCGGGATGCGGAGGCGCCCGAGGCGTCCCGGCGCTCGAGTCCGAGTACG	120	
Db	61	CGTTTGCCACCGGGATGCGGAGGCGCCCGAGGCGTCCCGGCGCTCGAGTCCGAGTACG	120	
Qy	121	ACCTCACGGCAAGTGATGCTTCTGGGAGACACAGCGCTCGCAAAACATGTTTCCTGA	180	
Db	121	ACCTCACGGCAAGTGATGCTTCTGGGAGACACAGCGCTCGCAAAACATGTTTCCTGA	180	
Qy	181	TCCAATTTAAAGACGGGGCTTCTCGCGGAACCTTCATAGCCACCGTCGGCATAGACT	240	
Db	181	TCCAATTTAAAGACGGGGCTTCTCGCGGAACCTTCATAGCCACCGTCGGCATAGACT	240	
Qy	241	TCAGGAACAAGTGTTGATGTTGGATGGCGTGAAGCTGAGATCTGGGACACCG	300	
Db	241	TCAGGAACAAGTGTTGATGTTGGATGGCGTGAAGCTGAGATCTGGGACACCG	300	

Db 241 TCAGGAACAAGGTGGTGACTGTGGATGGCGTGAGAGTGGAAGCTGGCAGATCTGGGACACCG 300
QY 301 CTGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCGCTTGC 360
Db 301 CTGGCAGGAACGGTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCGCTTGC 360
QY 361 TTCTGCTGTATGACATCACCAACAATCTTTCTTTCGACAAACATCAGGGCGCTGCTCAGCTG 420
Db 361 TTCTGCTGTATGACATCACCAACAATCTTTCTTTCGACAAACATCAGGGCGCTGCTCAGCTG 420
QY 421 AGATTATGATGATGCCACAGAGGAGTGTGATCATGCTGCTAGGCAACAAGCGCGGATA 480
Db 421 AGATTATGATGATGCCACAGAGGAGTGTGATCATGCTGCTAGGCAACAAGCGCGGATA 480
QY 481 TGAGCAGCAAGAGATGATCCGTTCCGAAGCAGGAGAGACCTTGGCCAGGAGTACGGGTG 540
Db 481 TGAGCAGCAAGAGATGATCCGTTCCGAAGCAGGAGAGACCTTGGCCAGGAGTACGGGTG 540
QY 541 TTCCCTTCTTGAGACAGGCGCAAGACTGGCATGGAATGAGGTGAGGTGAGCTTTCTGGCCA 600
Db 541 TTCCCTTCTTGAGACAGGCGCAAGACTGGCATGGAATGAGGTGAGGTGAGCTTTCTGGCCA 600
QY 601 TCGCCAAGGAACCTGAATACCGCGCGGCGATCAGCGGATGAGCCAGCTTCCAGATCC 660
Db 601 TCGCCAAGGAACCTGAATACCGCGCGGCGATCAGCGGATGAGCCAGCTTCCAGATCC 660
QY 661 GAGACTATGATAGTCCCAGGAAGCGCTCCAGCTGCTGCTCTTCATGTGAATCCCAG 720
Db 661 GAGACTATGATAGTCCCAGGAAGCGCTCCAGCTGCTGCTCTTCATGTGAATCCCAG 720
QY 721 GGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGAGCGCTTCCCGCTCCAGGCGCTGGC 780
Db 721 GGGCAGAGAGGAGGCTCTGGAGGCACACAGGATGAGCGCTTCCCGCTCCAGGCGCTGGC 780
QY 781 TTATTCCAGAGGCTGAGCCAATGGGAGAAAGATGGAGGACTCACTGACAGCGCGCTTC 840
Db 781 TTATTCCAGAGGCTGAGCCAATGGGAGAAAGATGGAGGACTCACTGACAGCGCGCTTC 840
QY 841 CTAGAGGGAGCTATCTCCAACTCCTACTGTAGTTTCTGCGGTCTCCCGCATCCACAG 900
Db 841 CTAGAGGGAGCTATCTCCAACTCCTACTGTAGTTTCTGCGGTCTCCCGCATCCACAG 900
QY 901 GGAGGTAAACACTTAGCTTTTATTATAGTACATATTTAATACCAAAAAGGCGC 960
Db 901 GGAGGTAAACACTTAGCTTTTATTATAGTACATATTTAATACCAAAAAGGCGC 960
QY 961 CTGGATCCCCAAAACCGAGGCTGGAGCTAGTGCCCTTTTGCTTTCTAGGACTTGGG 1020
Db 961 CTGGATCCCCAAAACCGAGGCTGGAGCTAGTGCCCTTTTGCTTTCTAGGACTTGGG 1020
QY 1021 GGGCGGCGCTCCTCTTAGCATACAAAGGTGTTGCTCCAGCTCAGGCGCCAGGGG 1080
Db 1021 GGGCGGCGCTCCTCTTAGCATACAAAGGTGTTGCTCCAGCTCAGGCGCCAGGGG 1080
QY 1081 ACACAGATGCACATTGGGGTGAGGCGAGTATGACTCCATCGCACCTCAGTTCAGCT 1140
Db 1081 ACACAGATGCACATTGGGGTGAGGCGAGTATGACTCCATCGCACCTCAGTTCAGCT 1140
QY 1141 GGACAGAGGCTCAGGTGACCCAGCGCTTCACTGTCTCCCGCTCCAGGAGCTTATCTTC 1200
Db 1141 GGACAGAGGCTCAGGTGACCCAGCGCTTCACTGTCTCCCGCTCCAGGAGCTTATCTTC 1200
QY 1201 GCGCCATCTCCCAATTAAGTGGCGCTTGTGCTGAGGAAGCAACAAAGCTCAGGGAAG 1260
Db 1201 GCGCCATCTCCCAATTAAGTGGCGCTTGTGCTGAGGAAGCAACAAAGCTCAGGGAAG 1260
QY 1261 ATAAGAGATGAGATGGAGGGGAGGACAGGCGCAGAGTAGGCTAGCTGGCT 1320
Db 1261 ATAAGAGATGAGATGGAGGGGAGGACAGGCGCAGAGTAGGCTAGCTGGCT 1320
QY 1321 ATCTCTGGCCTTACTAACACCCCGCTGGAGGCGATGCCCTTTTCTCCAGCACACAGCAC 1380
Db 1321 ATCTCTGGCCTTACTAACACCCCGCTGGAGGCGATGCCCTTTTCTCCAGCACACAGCAC 1380

QY 1381 ATTGGGCACTGGAATATTTGGTTCCAGGCTCCTGTCTCTCGACTTCAGATCCTCGGG 1440
Db 1381 ATTGGGCACTGGAATATTTGGTTCCAGGCTCCTGTCTCTCGACTTCAGATCCTCGGG 1440
QY 1441 GAGCCCTCCCCCGCTGAAATCCCTGGCTAGCTACCTTCCCTGCTGTGCACCTAAAAAC 1500
Db 1441 GAGCCCTCCCCCGCTGAAATCCCTGGCTAGCTACCTTCCCTGCTGTGCACCTAAAAAC 1500
QY 1501 CTCAGCTCAGAACTAGGAAAAGAGTTTGTGTTTATTTTGTGAAATGGAGTCTCGTCT 1560
Db 1501 CTCAGCTCAGAACTAGGAAAAGAGTTTGTGTTTATTTTGTGAAATGGAGTCTCGTCT 1560
QY 1561 GTCCCGCAGGCTCAGGTGCAATCTCCGCTCACTACAACCTCCACCTCCCTGGG 1620
Db 1561 GTCCCGCAGGCTCAGGTGCAATCTCCGCTCACTACAACCTCCACCTCCCTGGG 1620
QY 1621 GCTCAAGCGATCTCCACCTCAGCCGCCGAAGTAGTGGGACTATAGGTGTGTACCATC 1680
Db 1621 GCTCAAGCGATCTCCACCTCAGCCGCCGAAGTAGTGGGACTATAGGTGTGTACCATC 1680
QY 1681 ACACCTGGCTAATTTTGTATTTTGTAGACACAGGTTTCCGCATGTGCGCCAGGCTG 1740
Db 1681 ACACCTGGCTAATTTTGTATTTTGTAGACACAGGTTTCCGCATGTGCGCCAGGCTG 1740
QY 1741 GTCTTGAATTCCTGAGCTCAAGCAACCTGCGGCCCTCGGCCCTCCCAAGTACTGGGATTA 1800
Db 1741 GTCTTGAATTCCTGAGCTCAAGCAACCTGCGGCCCTCGGCCCTCCCAAGTACTGGGATTA 1800
QY 1801 CAGCGAAGAGGACCATGCCAGGCTAGATGTCTTTATCCCAATCCTTTGGCAGGAGTG 1860
Db 1801 CAGCGAAGAGGACCATGCCAGGCTAGATGTCTTTATCCCAATCCTTTGGCAGGAGTG 1860
QY 1861 CAGCTCCACAGGCGATTCTTCAAGCAGCTGGAAGTGTGTAGCCCTCCTGGGTTAAGAGCC 1920
Db 1861 CAGCTCCACAGGCGATTCTTCAAGCAGCTGGAAGTGTGTAGCCCTCCTGGGTTAAGAGCC 1920
QY 1921 AGATAAGGAAATCCCTTCTCTAGGTTTGGAAATGTTGTGAAAAAAGAGAAATCCC 1980
Db 1921 AGATAAGGAAATCCCTTCTCTAGGTTTGGAAATGTTGTGAAAAAAGAGAAATCCC 1980
QY 1981 TGCGCTCTGGAGCTGGGAGACAAGATTAAGCAACCTCCCTGCATGTATCCCTTT 2040
Db 1981 TGCGCTCTGGAGCTGGGAGACAAGATTAAGCAACCTCCCTGCATGTATCCCTTT 2040
QY 2041 GACCCCAAGCTCTGCTCCTCCTGACCAACCTGCTTCTTAACTTCTCAACAG 2100
Db 2041 GACCCCAAGCTCTGCTCCTCCTGACCAACCTGCTTCTTAACTTCTCAACAG 2100
QY 2101 ATACCAAGGCTTAACTGCTTACCTCCCTCTACTGAGTCAAGTTAGTGTGGGAGG 2160
Db 2101 ATACCAAGGCTTAACTGCTTACCTCCCTCTACTGAGTCAAGTTAGTGTGGGAGG 2160
QY 2161 TCACCCATTTCCAGTTAAACCAATGATGATAAACAAGTCATGTGGGTATGTC 2220
Db 2161 TCACCCATTTCCAGTTAAACCAATGATGATAAACAAGTCATGTGGGTATGTC 2220
QY 2221 TGGGTAGAGAGAGGGGTAGCAAGTTTCTGCTCCTTGGTGCATATCTCCCAAGC 2280
Db 2221 TGGGTAGAGAGAGGGGTAGCAAGTTTCTGCTCCTTGGTGCATATCTCCCAAGC 2280
QY 2281 TCTGATCCCTGCGATGGGAAGTGGACAGGAACATGAGGTCTAGCTCGAGGCATCTTT 2340
Db 2281 TCTGATCCCTGCGATGGGAAGTGGACAGGAACATGAGGTCTAGCTCGAGGCATCTTT 2340
QY 2341 ACTGCAAGCTCTCCGCTTGGAGGGGAGGGGAGGAGAGTAGTATGCGCTGCACATTT 2400
Db 2341 ACTGCAAGCTCTCCGCTTGGAGGGGAGGGGAGGAGAGTAGTATGCGCTGCACATTT 2400
QY 2401 CTGAGGCTACTGATTTGCTTTCAAGGCAGAAATCTGCTCTGAGCAGTCAGCGGCTCCA 2460
Db 2401 CTGAGGCTACTGATTTGCTTTCAAGGCAGAAATCTGCTCTGAGCAGTCAGCGGCTCCA 2460

QY 2461 GTTTGGGCCCCGATAGGAAGTTCTCCGTGGCTCCCTCAGGCAGACGAGGGAGGCTG 2520
Db 2461 GTTTGGGCCCCGATAGGAAGTTCTCCGTGGCTCCCTCAGGCAGACGAGGGAGGCTG 2520
QY 2521 ACATTGCCAGTCTCTTCTGGGGCCCCAAGCAGGTTGCAGGAGATCCAATCCATAGACAG 2580
Db 2521 ACATTGCCAGTCTCTTCTGGGGCCCCAAGCAGGTTGCAGGAGATCCAATCCATAGACAG 2580
QY 2581 CTCCTGGGCTCTTGCATTTGAGTTTTCAGAAATTAACCTGCAATTTTGGGAAACAAAA 2640
Db 2581 CTCCTGGGCTCTTGCATTTGAGTTTTCAGAAATTAACCTGCAATTTTGGGAAACAAAA 2640
QY 2641 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2674
Db 2641 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2674

RESULT 2
US-09-817-199A-3
; Sequence 3, Application US/09817199A
; Patent No. US20020142380A1
; GENERAL INFORMATION:
; APPLICANT: SHAO, Wei et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001187
; CURRENT APPLICATION NUMBER: US/09/817,199A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 13182
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(13182)
; OTHER INFORMATION: n = A,T,C or G
US-09-817-199A-3

Query Match 76.0%; Score 2031.8; DB 10; Length 13182;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2039; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 589 CTTTTCTGCCCATCCCAAGGAACCTGAATACCGGGCGGGCATCAGCGGATGAGCCCA 648
Db 10060 CCATTGTCTCTTCTTCAAGGGAACTGAATACCGGGCGGGCATCAGCGGATGAGCCCA 10119
QY 649 GCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAGCGCTCCAGCTGCTGCTCTCA 708
Db 10120 GCTTCCAGATCCGAGACTATGTAGAGTCCCAAGAGCGCTCCAGCTGCTGCTCTCA 10179
QY 709 TGTGAATCCAGGGGCGAGAGAGGCTCTGGAGGCACACAGGATGACGCTTCCCTC 768
Db 10180 TGTGAATCCAGGGGCGAGAGAGGCTCTGGAGGCACACAGGATGACGCTTCCCTC 10239
QY 769 CCCAGGCTGGCTTATTCACAGAGCTGAGCCAAATGGGGAGAAAGATGGAGACTCACTG 828
Db 10240 CCCAGGCTGGCTTATTCACAGAGCTGAGCCAAATGGGGAGAAAGATGGAGACTCACTG 10299
QY 829 CACAGCGCTTCTAGCAGGAGCTATCTCCAACTCTCTTGTAGTTCTCGGTCTCC 888
Db 10300 CACAGCGCTTCTAGCAGGAGCTATCTCCAACTCTCTTGTAGTTCTCGGTCTCC 10359
QY 889 CCGCATCCACAGGGAGGTAAACACTTAGCTTTTATTTAATAGTACATAATTAATAC 948
Db 10360 CCGCATCCACAGGGAGGTAAACACTTAGCTTTTATTTAATAGTACATAATTAATAC 10419
QY 949 CAAAAAGCGCTGGATCCCAAAAAACCGAGGCTGGAGCTAGTGGCCCTTTTGGCTTT 1008
Db 10420 CAAAAAGCGCTGGATCCCAAAAAACCGAGGCTGGAGCTAGTGGCCCTTTTGGCTTT 10479

QY 1009 CTAGGACTTGGGGGCGCGCCCTCCCTCCTAAGCATAAACAAAGTGGTGTGCTCCAGCT 1068
Db 10480 CTAGGACTTGGGGGCGCGCCCTCCCTCCTAAGCATAAACAAAGTGGTGTGCTCCAGCT 10539
QY 1069 CAGCCCCAGGGGACACAGATGCACCTTTGGGGGTGAGGGCAGGTAAATGACTCCATCCGACC 1128
Db 10540 CAGCCCCAGGGGACACAGATGCACCTTTGGGGGTGAGGGCAGGTAAATGACTCCATCCGACC 10599
QY 1129 CTCAGTTGAGCTGGACAGAGGCTCAGGTGACCCAGCCCTTCACTGCTCCCGCTCTCCAG 1188
Db 10600 CTCAGTTGAGCTGGACAGAGGCTCAGGTGACCCAGCCCTTCACTGCTCCCGCTCTCCAG 10659
QY 1189 GAGCTTATCTTCGCCCCATCTCCCAATAAGTGGGCCCTTGTGCTGTGAGGAGACAAA 1248
Db 10660 GAGCTTATCTTCGCCCCATCTCCCAATAAGTGGGCCCTTGTGCTGTGAGGAGACAAA 10719
QY 1249 GCCTCAGGGAAGATAGAGATATGGAGATGGAGGGGAGGACAGGGGACAGAGTAGG 1308
Db 10720 GCCTCAGGGAAGATAGAGATATGGAGATGGAGGGGAGGACAGGGGACAGAGTAGG 10779
QY 1309 GTCTAGCTGGCTATCTCGGCCCTTACTAACACCCCTCGAGGSCATGCCCTTTCTCCA 1368
Db 10780 GTCTAGCTGGCTATCTCGGCCCTTACTAACACCCCTCGAGGSCATGCCCTTTCTCCA 10839
QY 1369 GCACACAGCACAATTGGGGCAGCTGGAATATTGTTCCAGGCTCCTGTTCTCTGGACTT 1428
Db 10840 GCACACAGCACAATTGGGGCAGCTGGAATATTGTTCCAGGCTCCTGTTCTCTGGACTT 10899
QY 1429 CAGATCTGGGGAGGCCCTCCCCCTTGAATCCCTGGCTTACCTTCCCTGCTGT 1488
Db 10900 CAGATCTGGGGAGGCCCTCCCCCTTGAATCCCTGGCTTACCTTCCCTGCTGT 10959
QY 1489 GCACCTAATAAACCCTCAGGTCAGAACTAGAGAAAGAGTTTGTATTTTGTGAAATG 1548
Db 10960 GCACCTAATAAACCCTCAGGTCAGAACTAGAGAAAGAGTTTGTATTTTGTGAAATG 11019
QY 1549 GAGTCTGCTTCTGTCGCCAGGCTCAGGTGCAAGTGTGCAATCTCCGCTCACTACAACT 1608
Db 11020 GAGTCTGCTTCTGTCGCCAGGCTCAGGTGCAAGTGTGCAATCTCCGCTCACTACAACT 11079
QY 1609 CCACCTCTGGGGCTCAAGCGATCTCCACCTCAGCCCGCAGTAGCTGGGACTATAG 1668
Db 11080 CCACCTCTGGGGCTCAAGCGATCTCCACCTCAGCCCGCAGTAGCTGGGACTATAG 11139
QY 1669 GTGTGTACCATCACACTGSCCTAATTTTGTATTTTGTAGACACAGGTTTCCGCATG 1728
Db 11140 GTGTGTACCATCACACTGSCCTAATTTTGTATTTTGTAGACACAGGTTTCCGCATG 11199
QY 1729 TTGCCCAGGCTGGTCTTGAATTTCTGAGCTCAAGCAACCTGCGGCTCGGCTCCCAAA 1788
Db 11200 TTGCCCAGGCTGGTCTTGAATTTCTGAGCTCAAGCAACCTGCGGCTCGGCTCCCAAA 11259
QY 1789 GTACTGGGATACACGCAAGGACCATGCCCAGGCTAGATGTGCTTATCCCAATCCT 1848
Db 11260 GTACTGGGATACACGCAAGGACCATGCCCAGGCTAGATGTGCTTATCCCAATCCT 11319
QY 1849 TTGGCAGCATCGCTCCACAGGCTTCTTCAAGCAGCTCAAGTGTGTTAGCCCTCCT 1908
Db 11320 TTGGCAGCATCGCTCCACAGGCTTCTTCAAGCAGCTCAAGTGTGTTAGCCCTCCT 11379
QY 1909 GGGTTAAAGCCAGATAGAGAAATCCCTTTCTAGGTTTGGAAATGTGTGTGAAAAAA 1968
Db 11380 GGGTTAAAGCCAGATAGAGAAATCCCTTTCTAGGTTTGGAAATGTGTGTGAAAAAA 11439
QY 1969 AAGAGAAATCCCTGGCTCTGGAGCTGGGGAGACAGATTAAAGCAAACTCCCTGAC 2028
Db 11440 AAGAGAAATCCCTGGCTCTGGAGCTGGGGAGACAGATTAAAGCAAACTCCCTGAC 11499
QY 2029 ATGTATCCCTTTGACCCCAAGCTCTGCCCTCCCTGACCCCATGCCCTTCCCTTAA 2088
Db 11500 ATGTATCCCTTTGACCCCAAGCTCTGCCCTCCCTGACCCCATGCCCTTCCCTTAA 11559
QY 2089 CTTCTCAACAGATACAGGGGCTTAACTGCTTTTACCTCCCTCTCTACTGAGTCAGTTA 2148

```
Db 11560 CTTCTCAACAGATACGAGGCTAAACCTGCTTTACTCCCTCCCTACTGAGTCAGGTTA 11619
QY 2149 GGTGGTGGGAGGTCAACCCATTTCCGAGTTAAACCAATGATATAGTAAACAAAGTCA 2208
Db 11620 GGTGGTGGGAGGTCAACCCATTTCCGAGTTAAACCAATGATATAGTAAACAAAGTCA 11679
QY 2209 TGTGGTATGCTGGGGTAGAGAGGGGTAGCAAGTTTCATGTGCTCCTCTTGGTCACAT 2268
Db 11680 TGTGGTATGCTGGGGTAGAGAGGGGTAGCAAGTTTCATGTGCTCCTCTTGGTCACAT 11739
QY 2269 ATCTCCCAAGCTCTCATCCCTGCCATGGAAGTGACAGGAAACATGAGGTGATGACCT 2328
Db 11740 ATCTCCCAAGCTCTCATCCCTGCCATGGAAGTGACAGGAAACATGAGGTGATGACCT 11799
QY 2329 GCAGGCACTTTACTGCAGCTCTGCCGGCTGGAGGGGAGAGGGGAGGAAAGTATG 2388
Db 11800 GCAGGCACTTTACTGCAGCTCTGCCGGCTGGAGGGGAGAGGGGAGGAAAGTATG 11859
QY 2389 CGTGCACATTTCTGAGGCTACTGCAATTTGCTTTCAAGCAGAAATCTTGCTCTGAGCAG 2448
Db 11860 CGTGCACATTTCTGAGGCTACTGCAATTTGCTTTCAAGCAGAAATCTTGCTCTGAGCAG 11919
QY 2449 TCAGCGCTCCAGTTTGGGCCGATAGGAAGTTCTCCGTGGCTCCCTCAGSCAGCA 2508
Db 11920 TCAGCGCTCCAGTTTGGGCCGATAGGAAGTTCTCCGTGGCTCCCTCAGSCAGCA 11979
QY 2509 GGGAGGAGGCTGACATTTGCCAGTCTCTCTGGGGCCCAAGGCAGGTTGCAGGAGATCCAA 2568
Db 11980 GGGAGGAGGCTGACATTTGCCAGTCTCTCTGGGGCCCAAGGCAGGTTGCAGGAGATCCAA 12039
QY 2569 TCCCATAGACAGCTCGGGCTCTTGCATTTGAGTTTTCAGAACTGAGATTTT 2628
Db 12040 TCCCATAGACAGCTCGGGCTCTTGCATTTGAGTTTTCAGAACTGAGATTTT 12099
QY 2629 TGGAAAGCAA 2639
Db 12100 TGGAAAGCACA 12110

RESULT 3
US-09-794-257-13
; Sequence 13, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: Human G-Proteins
; FILE REFERENCE: 35800/209285
; CURRENT APPLICATION NUMBER: US/09/794,257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (124)...(699)
US-09-794-257-13

Query Match 40.6%; Score 1086; DB 10; Length 1116;
Best Local Similarity 99.5%; Pred. No. 1.6e-181;
Matches 1089; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 37 GGGACATGAGGGCAGCCAGGCGCGTTGCCACCGGGATGCGAGGCCCGCGAGCGGT 96
Db 23 GCGTCGGGAGGGCAGCCAGGCGCGTTGCCACCGGGATGCGAGGCCCGCGAGCGGT 82
QY 97 CCCCCTCCAGTCGAGGCTAGACCTCAGGGGCAAGGTGATGCTCTCTGGGAGACACAG 156
```

```
Db 83 CCCCCTCCAGTCGAGTCGAGCTCAGGCTCAGGGCAAGGTGATGCTCTCTGGGAGACACAG 142
QY 157 GGTGCGGCAAAACATGTTTCTCTGATCCAAATTCAGAGACGGGCTTCTCTGTCGGGAACCT 216
Db 143 GGTGCGGCAAAACATGTTTCTCTGATCCAAATTCAGAGACGGGCTTCTCTGTCGGGAACCT 202
QY 217 TCATAGCCACCGTCGCGCATAGACTTCAGGAACAAGTGTGACTGTGGATGGCGTGAGAG 276
Db 203 TCATAGCCACCGTCGCGCATAGACTTCAGGAACAAGTGTGACTGTGGATGGCGTGAGAG 262
QY 277 TGAAGTCGAGACTCTGGGACACCGCTGGGCAAGCGTTCCGAAGCGTCAACCATGCTT 336
Db 263 TGAAGTCGAGACTCTGGGACACCGCTGGGCAAGCGTTCCGAAGCGTCAACCATGCTT 322
QY 337 ATTACAGAGATGCTCAGGCTTGTCTGTGTATGACATCACCAAAATCTTCTTTTCG 396
Db 323 ATTACAGAGATGCTCAGGCTTGTCTGTGTATGACATCACCAAAATCTTCTTTTCG 382
QY 397 ACAACATCAGGCTGGCTGCTGACTGAGATTCATGAGTATGCCAGAGGAGGTGGTGATCA 456
Db 383 ACAACATCAGGCTGGCTGCTGACTGAGATTCATGAGTATGCCAGAGGAGGTGGTGATCA 442
QY 457 TGTGCTAGGCAACAGGCGGATATGAGCAGGAGAAAGTGTATCCCTTCCGAAGACGGAG 516
Db 443 TGTGCTAGGCAACAGGCGGATATGAGCAGGAGAAAGTGTATCCCTTCCGAAGACGGAG 502
QY 517 AGACCTTGCCAGGAGTACGCTTGTCTCTGTGGAGCAGCGCCAAAGACTGGCATGA 576
Db 503 AGACCTTGCCAGGAGTACGCTTGTCTCTGTGGAGCAGCGCCAAAGACTGGCATGA 562
QY 577 ATGTGAGTGTAGCTTCTGGGCTATGCCAAGAACTGAAATACCGGGCGGGGCATCAGG 636
Db 563 ATGTGAGTGTAGCTTCTGGGCTATGCCAAGAACTGAAATACCGGGCGGGGCATCAGG 622
QY 637 CGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGTCCCAAGAGAGAGCGCTCCAGCT 696
Db 623 CGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGTCCCAAGAGAGAGCGCTCCAGCT 682
QY 697 GCTGCTCTTCAATGTAATCCAGGGGCGAGAGAGGCTCTTGGAGGCACACAGGATGC 756
Db 683 GCTGCTCTTCAATGTAATCCAGGGGCGAGAGAGGCTCTTGGAGGCACACAGGATGC 742
QY 757 AGCTTCCCTCCAGGCTGCTTATTCGAAGAGGCTGAGCCATGGGAGAAAGATG 816
Db 743 AGCTTCCCTCCAGGCTGCTTATTCGAAGAGGCTGAGCCATGGGAGAAAGATG 802
QY 817 GAGGACTCACTGCACAGCGCTTCTTAGCAGGAGGCTATATCTCAACTCTCTACTTGAGTT 876
Db 803 GAGGACTCACTGCACAGCGCTTCTTAGCAGGAGGCTATATCTCAACTCTCTACTTGAGTT 862
QY 877 CCTGCGGTCTCCCGCATCCACAGGAGGTTAAACACTTAGCTTTTATTTAATAGTAC 936
Db 863 CCTGCGGTCTCCCGCATCCACAGGAGGTTAAACACTTAGCTTTTATTTAATAGTAC 922
QY 937 ATAAATTAACAAAAAGGCGCTTGATCCCAAAAAACCGAGGCTGGGAGCTAGTG 996
Db 923 ATAAATTAACAAAAAGGCGCTTGATCCCAAAAAACCGAGGCTGGGAGCTAGTG 982
QY 997 CCCTTTTGTCTTAGGACTTGGGGGCGGCGCTCCCTCTAAGCATACCAAGGTGGT 1056
Db 983 CCCTTTTGTCTTAGGACTTGGGGGCGGCGCTCCCTCTAAGCATACCAAGGTGGT 1042
QY 1057 GTTGCTCCAGCTCAGCCCGGAGGAGACACAGATGCACCTTTTGGGGGTGAGGSCAGGTAATGA 1116
Db 1043 GTTGCTCCAGCTCAGCCCGGAGGAGACACAGATGCACCTTTTGGGGGTGAGGSCAGGTAATGA 1102
QY 1117 CTCCTATCGACCT 1130
Db 1103 CTCCTATCGACCT 1116

RESULT 4
```

Query Match	32.6%	Score 873	DB 12	Length 875
Best Local Similarity	100.0%	Pred. No. 2.4e-144		
Matches 873	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	46	CGSGCAGCCAGGCGCGTTCACACCCGGGATGCGAGGCCGCCGAGCGCTCCCGCGCCT	105	
DB	1	CGSGCAGCCAGGCGCGTTCACACCCGGGATGCGAGGCCGCCGAGCGCTCCCGCGCCT	60	
QY	106	GCAGTCCGAGCTACGACCTCACGGGCAAGTGATGCTTCCTGGGAGACACAGGGCGTCGGCA	165	
DB	61	GCAGTCCGAGCTACGACCTCACGGGCAAGTGATGCTTCCTGGGAGACACAGGGCGTCGGCA	120	
QY	166	AAACATGTTTCTGATCCAATTCAAACACGGGGGCTTCCTGTCGGGAACCTTCATAGCCA	225	
DB	121	AAACATGTTTCTGATCCAATTCAAACACGGGGGCTTCCTGTCGGGAACCTTCATAGCCA	180	
QY	226	CCGTGGCATAGACTTCAGGAACAAGTGTGTACTGTGGATGGCGTGAGAGTGAAGCTGC	285	
DB	181	CCGTGGCATAGACTTCAGGAACAAGTGTGTACTGTGGATGGCGTGAGAGTGAAGCTGC	240	

Query Match	25.6%	Score 683.4	DB 10	Length 24707
Best Local Similarity	99.9%	Pred. No. 4.9e-111		
Matches 684	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	1955	GTGTTGTGAAAAAAGAGAAATCCTCGCTCCTCGAGCTGTGGGAGACAAGATTAAAGC	2014	
Db	1	GTGTTGTGAAAAAAGAGAAATCCTCGCTCCTCGAGCTGTGGGAGACAAGATTAAAGC	60	

QY 2015 AAACCTCCCTGACATGATATCCCTTTGACCCCAAGCTCTGCCCTCCCTGACCAACCCAT 2074
Db 61 AAACCTCCCTGACATGATATCCCTTTGACCCCAAGCTCTGCCCTCCCTGACCAACCCAT 120
QY 2075 GCCCTTTCCTTTAACTTCTCAACAGATACAGGCGCTAAACTGCTTTACTCTCCCTCCCT 2134
Db 121 GCCCTTTCCTTTAACTTCTCAACAGATACAGGCGCTAAACTGCTTTACTCTCCCTCCCT 180
QY 2135 ACTGAGTCAGGTAGGTGGGAGGTCAACCATTTCCGAGTTAAACCAATGCAATATGA 2194
Db 181 ACTGAGTCAGGTAGGTGGGAGGTCAACCATTTCCGAGTTAAACCAATGCAATATGA 240
QY 2195 GTAACAAAGTATGTTGGGTATGCTGGGGTAGAGAGGGGTAGCAAGTTCATGTGTC 2254
Db 241 GTAACAAAGTATGTTGGGTATGCTGGGGTAGAGAGGGGTAGCAAGTTCATGTGTC 300
QY 2255 CTCTTGGTCACATATCTCCCAAGCTCTGATCCCTCCCATGGAGTGGACAGGAACA 2314
Db 301 CTCTTGGTCACATATCTCCCAAGCTCTGATCCCTCCCATGGAGTGGACAGGAACA 360
QY 2315 TGAGGTGATGACCTGACGCGATCTTTACTGACGCTCTGCCGCTCTGGAGGGGAGAGGGG 2374
Db 361 TGAGGTGATGACCTGACGCGATCTTTACTGACGCTCTGCCGCTCTGGAGGGGAGAGGGG 420
QY 2375 GAGGAAGAAGTATGCGGTGACATTTCTGAGGCTACTGCAATTTGCTTTCAAGGCAGAAAT 2434
Db 421 GAGGAAGAAGTATGCGGTGACATTTCTGAGGCTACTGCAATTTGCTTTCAAGGCAGAAAT 480
QY 2435 CTGCTGAGCAGTACAGCGCTCCAGTTTGGCCGATAGGAAGTCTCCGTCGGCCTC 2494
Db 481 CTGCTGAGCAGTACAGCGCTCCAGTTTGGCCGATAGGAAGTCTCCGTCGGCCTC 540
QY 2495 CCTCAGCAGCAGGAGGAGGCTGACATTTGCCAGTCTCTTCTGGGGCCCAAGGCAGGT 2554
Db 541 CCTCAGCAGCAGGAGGAGGCTGACATTTGCCAGTCTCTTCTGGGGCCCAAGGCAGGT 600
QY 2555 TGAGGAGATCCCAATCCATAGACAGCTCTGGGCTCTTGCATTTGAGTTTTCAGAAAT 2614
Db 601 TGAGGAGATCCCAATCCATAGACAGCTCTGGGCTCTTGCATTTGAGTTTTCAGAAAT 660
QY 2615 AAACCTGAGTATTTGGAAAGCAA 2639
Db 661 AAACCTGAGTATTTGGAAAGCACA 685

RESULT 6
US-09-817-199a-24
; Sequence 24, Application US/09817199a
; Patent No. US20020142380A1
; GENERAL INFORMATION:
; APPLICANT: SHAO, Wei et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO01187
; CURRENT APPLICATION NUMBER: US/09/817,199a
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-817-199a-24

Query Match 22.5%; Score 600.6; DB 10; Length 601;
Best Local Similarity 99.8%; Pred. No. 8.6e-97;
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1983 GCTCCTGGAGCTGGTGGGAGACAGATTAAGCAAACTCCCTGACATGATCCCTTTGA 2042
Db 1 GCTCCTGGAGCTGGTGGGAGACAGATTAAGCAAACTCCCTGACATGATCCCTTTGA 60

..

QY 2043 CCCAAGCTCTGCTCCTCCTGACACCCATGCCCTTTCTTAACTTCTCAAAACAGAT 2102
Db 61 CCCAAGCTCTGCTCCTCCTGACACCCATGCCCTTTCTTAACTTCTCAAAACAGAT 120
QY 2103 ACCAGGSCCTAAACTGCTTTACCTCCCTCCTACTAGTCAGGTAGGTGGGAGGTC 2162
Db 121 ACCAGGSCCTAAACTGCTTTACCTCCCTCCTACTAGTCAGGTAGGTGGGAGGTC 180
QY 2163 ACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAACCAAGTCATGTGGTATGTCG 2222
Db 181 ACCCATTTCCGAGTTAAACCAATGCAATATGAGTAAACCAAGTCATGTGGTATGTCG 240
QY 2223 GGTAGAGAGAGGGGTAGCAAGTTCATGTCTCCTTGGTCAATATCCTCCAAAGCTC 2282
Db 241 GGTAGAGAGAGGGGTAGCAAGTTCATGTCTCCTTGGTCAATATCCTCCAAAGCTC 300
QY 2283 TGATCCCTGCCATGGGAGTGGACAGAAACATGAGTCATGACCTCGAGGCATCTTTAC 2342
Db 301 YGATCCCTGCCATGGGAGTGGACAGAAACATGAGTCATGACCTCGAGGCATCTTTAC 360
QY 2343 TGCAGCTCTGCCGCTCTGGAGGGGAGGGGAGGAGTATGCGCTGCACATTTCT 2402
Db 361 TGCAGCTCTGCCGCTCTGGAGGGGAGGGGAGGAGTATGCGCTGCACATTTCT 420
QY 2403 GAGGCTACTGCAATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAGTCAGCGGCTCCAGT 2462
Db 421 GAGGCTACTGCAATTTGCTTTCAAGGCAGAAATCTTGCTCTGAGCAGTCAGCGGCTCCAGT 480
QY 2463 TTGGGCCCCGATAGGAGTTCCTCGTGGCTCCTCCTCAGGACAGCAGGAGGAGGCTGAC 2522
Db 481 TTGGGCCCCGATAGGAGTTCCTCGTGGCTCCTCCTCAGGACAGCAGGAGGAGGCTGAC 540
QY 2523 ATTGCCAGTCTCTTCTGGGCCCCAAGGCAGGTTGCAGGAGATCCAATCCCATAGACAGCT 2582
Db 541 ATTGCCAGTCTCTTCTGGGCCCCAAGGCAGGTTGCAGGAGATCCAATCCCATAGACAGCT 600
QY 2583 C 2583
Db 601 C 601

RESULT 7
US-09-794-257-15
; Sequence 15, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1
; FILE REFERENCE: Human G-Proteins
; CURRENT APPLICATION NUMBER: US/09/794,257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 576
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-794-257-15

Query Match 21.5%; Score 576; DB 10; Length 576;
Best Local Similarity 100.0%; Pred. No. 1.7e-92;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 ATGCTTCTGGAGACACAGCGCTCGGCAAAACATGTTTCTGATCCAAATCAAGACGGG 197
Db 1 ATGCTTCTGGAGACACAGCGCTCGGCAAAACATGTTTCTGATCCAAATCAAGACGGG 60
QY 198 GCCTTCTGTCGGAACTTCATAGCCACCGTCGGGATAGACTTCAGGAACAAGTGCGT 257
Db 61 GCCTTCTGTCGGAACTTCATAGCCACCGTCGGGATAGACTTCAGGAACAAGTGCGT 120

QY 258 ACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGCAGGAACGGTTC 317
DB 121 ACTGTGGATGGCGTGAGAGTGAAGCTGCAGATCTGGGACACCGCTGGCAGGAACGGTTC 180
QY 318 CGAAGCGTCACCCATGCTTTATACAGAGATGCTCAGGCTTGGCTTCTGCTGTATGACATC 377
DB 181 CGAAGCGTCACCCATGCTTTATACAGAGATGCTCAGGCTTGGCTTCTGCTGTATGACATC 240
QY 378 ACCAACAATCTCTTTTCGACAACATCAGGGCTGGCTCAGTCACTGAGATTCAATGATGCC 437
DB 241 ACCAACAATCTCTTTTCGACAACATCAGGGCTGGCTCAGTCACTGAGATTCAATGATGCC 300
QY 438 CAGAGGAGCGTGGTATCATCTGCTAGGCAACAAAGGCGGATATGAGCAGGCAAAAGAGTG 497
DB 301 CAGAGGAGCGTGGTATCATCTGCTAGGCAACAAAGGCGGATATGAGCAGGCAAAAGAGTG 360
QY 498 ATCCGTTCCGAGAGCGGAGACCTTGGCCAGGGAGTACGGTGTTCCTTCTGAGAGACC 557
DB 361 ATCCGTTCCGAGAGCGGAGACCTTGGCCAGGGAGTACGGTGTTCCTTCTGAGAGACC 420
QY 558 AGCGCAAGACTGGCATCAATGTGGAGTACCTTTCTGGCCATCGCCCAAGGAACCTGAAA 617
DB 421 AGCGCAAGACTGGCATCAATGTGGAGTACCTTTCTGGCCATCGCCCAAGGAACCTGAAA 480
QY 618 TACCGGGCGGCGATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 677
DB 481 TACCGGGCGGCGATCAGCGGATGAGCCAGCTTCCAGATCCGAGACTATGTAGAGTCC 540
QY 678 CAGAAGAAGCGCTCAGCTGCTGCTTCATGTGA 713
DB 541 CAGAAGAAGCGCTCAGCTGCTGCTTCATGTGA 576

RESULT 8

US-09-817-199A-25
; Sequence 25, Application US/09817199A
; Patent No. US20020142380A1
; GENERAL INFORMATION:
; APPLICANT: SHAO, Wei et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO01187
; CURRENT APPLICATION NUMBER: US/09/817,199A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-817-199A-25

Query Match 21.4%; Score 573; DB 10; Length 601;
Best Local Similarity 99.7%; Pred. No. 5.7e-92;
Matches 573; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2055 GACCAACCATGCCCTTCCCTTTAACTTCTCAACAGATACACAGGCGCTTAACTGCTTAC 2124
DB 1 GACCAACCATGCCCTTCCCTTTAACTTCTCAACAGATACACAGGCGCTTAACTGCTTAC 60
QY 2125 CTCCTCCTACTAGTCAGGTAGGTGGTGGAGGTACCCATTTCCGAGTTAAACCAA 2184
DB 61 CTCCTCCTACTAGTCAGGTAGGTGGTGGAGGTACCCATTTCCGAGTTAAACCAA 120
QY 2185 TGCATATGATTAACAAGTCAATGTGGGTATGCTGGGTAGAGAGGGGTAGCAAG 2244
DB 121 TGCATATGATTAACAAGTCAATGTGGGTATGCTGGGTAGAGAGGGGTAGCAAG 180
QY 2245 TTCATGTCTCTCTGCTGCTACATATCTCCAAAGCTCTGATCCCTGCGATGGGAAGTGG 2304
DB 181 TTCATGTCTCTCTGCTGCTACATATCTCCAAAGCTCTGATCCCTGCGATGGGAAGTGG 240

QY 2305 ACAGGAACATCAGCTCATGCTGAGGCATCTTTACTGCAGCTCTGCGCGCTGGAGG 2364
DB 241 ACAGGAACATCAGCTCATGCTGAGGCATCTTTACTGCAGCTCTGCGCGCTGGAGG 300
QY 2365 GGGAGAGGGGAGGAGAGATATGCGTGCACATTTCTGAGGCTACTGCAATTTGCTTTCA 2424
DB 301 RGGAGAGGGGAGGAGAGATATGCGTGCACATTTCTGAGGCTACTGCAATTTGCTTTCA 360
QY 2425 AGGCAGAAATCTTCTCTGAGCAGTCAGCGCTCCAGTTTGGGCCGATAAGAAAGTTCT 2484
DB 361 AGGCAGAAATCTTCTCTGAGCAGTCAGCGCTCCAGTTTGGGCCGATAAGAAAGTTCT 420
QY 2485 CCGTGGCCTCCCTCAGGAGAGCAGGAGGCTGACATTTGCCAGTCTCTTCTTGGGGCC 2544
DB 421 CCGTGGCCTCCCTCAGGAGAGCAGGAGGCTGACATTTGCCAGTCTCTTCTTGGGGCC 480
QY 2545 CAAGCAGGTTGCAGAGATCCAAATCCCATAGACAGCTCTTGGGCTCTTTCGATTTGAGTT 2604
DB 481 CAAGCAGGTTGCAGAGATCCAAATCCCATAGACAGCTCTTGGGCTCTTTCGATTTGAGTT 540
QY 2605 TTTCAGAAATTAACATGCACTATTTTGGAAAGCAA 2639
DB 541 TTTCAGAAATTAACATGCACTATTTTGGAAAGCAA 575

RESULT 9

US-09-867-550-1811
; Sequence 1811, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela,
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: NO. US20020082206A1el Polynucleotides from Atherogenic Cells a
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1811
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: Wherein n is one of a or t or c or g
US-09-867-550-1811

Query Match 16.6%; Score 443.4; DB 10; Length 447;
Best Local Similarity 99.8%; Pred. No. 2.3e-69;
Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 GCGCGGCACTGCTCACCTCTCTGTCAGGAGACATGACGGCAGCCAGGCGCGCTTGGCCAC 70
DB 3 GCGCGGCACTGCTCACCTCTCTGTCAGGAGACATGACGGCAGCCAGGCGCGCTTGGCCAC 62
QY 71 CCGGATGCGAGGCGCGCGGAGCGCTCCCGGCTTCAGTCCGAGTCCGAGCTAGCAGCTCACGGG 130
DB 63 CCGGATGCGAGGCGCGCGGAGCGCTCCCGGCTTCAGTCCGAGTCCGAGCTAGCAGCTCACGGG 122
QY 131 CAAGTGTATGCTTCTGGGAGACACAGGCTCGGCAAAACATCTTCCCTGATCCAATTCAA 190
DB 123 CAAGTGTATGCTTCTGGGAGACACAGGCTCGGCAAAACATCTTCCCTGATCCAATTCAA 182
QY 191 AGACGGGCGCTTCTCTGTCGGAACCTTCATAGCCACCGCTCGGCATAGACTTCAGGAACAA 250

Db 183 AGAGGGGGCTTCTGTCCGGAACCTTCATAGCCACCGTGGCATAGACTTCAGGAACAA 242

QY 251 GGTGGTGACTGTGGATGGCTGAGATGAAGCTGCAGATCTGGGACACGCTGGGCAGGA 310

Db 243 GGTGGTGACTGTGGATGGCTGAGATGAAGCTGCAGATCTGGGACACGCTGGGCAGGA 302

QY 311 ACGTTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCCTTGCTTCTGCTGA 370

Db 303 ACGTTTCCGAAGCGTCACCCATGCTTATTACAGAGATGCTCAGGCCCTTGCTTCTGCTGA 362

QY 371 TGACATCACCACAACAAATCTTTCACACACATCAGGGGCTGCTCAGTGAATTCATGA 430

Db 363 TGACATCACCACAACAAATCTTTCACACACATCAGGGGCTGCTCAGTGAATTCATGA 422

QY 431 GTATGCCAGAGGAGCTGGTGATC 455

Db 423 GTATGCCAGAGGAGCTGGTGATC 447

RESULT 10

US-09-764-877-2275/c

; Sequence 2275, Application US/09764877

; Patent No. US20020147140A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC005

; CURRENT APPLICATION NUMBER: US/09/764,877

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 4031

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2275

; LENGTH: 21833

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-877-2275

Query Match 7.5%; Score 200.6; DB 10; Length 21833;

Best Local Similarity 74.1%; Pred. No. 9.9e-27;

Matches 254; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 1525 TTTTGTGTTTTATTTTTTGAATGAGTCTCGTTCGTGCGCCAGGCTGAGTGCAGTAG 1584

Db 5032 TTTTGTGTTTTATTTTTTGAATGAGTCTCGTTCGTGCGCCAGGCTGAGTGCAGTAG 1584

QY 1585 TGAATCTCGGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCTCCACCTCAG 1644

Db 4972 TGAATCTCGGCTCACTACAACCTCCACTCCCTGGGGCTCAAGCGATCTCCACCTCAG 1644

QY 1645 CCGCCGAAGTAGCTGGGACTATAGTGTGTACCATCACACCTGGGTAATTTTGTATTTT 1704

Db 4912 CCGCCGAAGTAGCTGGGACTATAGTGTGTACCATCACACCTGGGTAATTTTGTATTTT 1704

QY 1705 TTGTAGACACAGGTTTCCCATGTTGCCAGGCTGCTTGAATTCGTAGCTCAAGCA 1764

Db 4852 TAGTAGACACAGGTTTCCCATGTTGCCAGGCTGCTTGAATTCGTAGCTCAAGCA 1764

QY 1765 ACTGCGCGGCTCGGCTCCCAAGTACTGGGATGATGAGGACAGAGGACCATGCCAGG 1824

Db 4792 ATCCACCTGCTTATCCCAATCTTTGGCAGGCTGAGGATACAGGCTGAGCCGCTCC 4733

QY 1825 CTAGATGTGCTTATCCCAATCTTTGGCAGGCTGAGGATACAGGCTGAGCCGCTCC 1867

Db 4732 CCGTGAATTTATTCATGATATTTCTTGATGAAAGGCCCTCC 4690

RESULT 11

US-09-764-877-3373/c

; Sequence 3373, Application US/09764877

; Patent No. US20020147140A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC005

; CURRENT APPLICATION NUMBER: US/09/764,877

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 4031

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3373

; LENGTH: 1130

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-877-3373

Query Match 7.5%; Score 199.8; DB 10; Length 1130;

Best Local Similarity 77.1%; Pred. No. 9.2e-27;

Matches 243; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1516 GGAAAAGAGTTTTTGTGTTTTTAAATGGAGTCTCGTTCGTGCGCCAGGCTGAG 1575

Db 673 GTAATTAATTTTCCCTTTTTTTTGGAGACAGGSCCTAGCTCTGACACCCAGGCTGAA 614

QY 1576 GTGCAGTAGTGAATCTCGGCTCACTACAACCTCCACTCCCTGGGGCTCAAGGATCCTC 1635

Db 613 GTGCAGTGGTGTGATCTCAGCTCACTGCAACCTCTGCTTCTGGACTCAAGTGATCCTC 554

QY 1636 CCACCTCAGCGCCGGAAGTAGCTGGGACTATAGGTGTGTACCATCACACCTGCTTAATTT 1695

Db 553 CCACCTCAGCGCTTAAAGTAGCTGGGACTTACAAATGCACACCCACACCTGGTTAATTT 494

QY 1696 TTGTATTTTTGTAGACACAGGTTTTGCGCATGTTGCCAGGCTGGTCTTGAATTCCTGA 1755

Db 493 TTTGATTTTTTGTAGACACAGAGTTTTTGCACGTTGCCAGGCTGGTCTGGAACCTCTCG 434

QY 1756 GCTCAAGCAACCTGCGGGCTCGGCTCCCAAGTACTGGGATGTACACCCAGGACACC 1815

Db 433 GCTCAAGCGATCCGCCACCTTGGCTCCCAAAGTCTAGGATTACAGGAGTGAGCCACA 374

QY 1816 ATGCCCGAGGCTAGAT 1830

Db 373 ATGCCCTGGCCGCAAT 359

RESULT 12

US-09-764-877-3236

; Sequence 3236, Application US/09764877

; Patent No. US20020147140A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC005

; CURRENT APPLICATION NUMBER: US/09/764,877

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 4031

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3236

; LENGTH: 15745

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-877-3236

Query Match 7.4%; Score 198.6; DB 10; Length 15745;

Best Local Similarity 76.7%; Pred. No. 2.1e-26;

Matches 243; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1514 TAGGAAAAGAGTTTTTGTGTTTTTAAATGGAGTCTCGTTCGTGCGCCAGGCTG 1573

Db 6486 TAAAGTTACTAATTTTTTTTTTTTTTTTGGAGACAGGATCTGCTCTGTCACCCAGGCTG 6545

QY 1574 AGGTGCAGTAGTGAATCTCGGCTCACTACACCTCCACTCCCTGGGGCTCAAGGATCC 1633

Db 6546 GAGTCAGTGGTACGATCTCAGCTCACTGTAGCCCTTAACCCACCAGGCTTATGCGCTCC 6605

Query Match	7.4%	Score 198.6;	DB 10;	Length 17217;
Best Local Similarity	76.3%	Pred. No. 2.1le-36;		
Matches 271; Conservative	0;	Mismatches 79;	Indels 5;	Gaps 2;

; FILE REFERENCE: PC005

```

Db 2799 ACCTCAGCTTCCCAAGTAACCTGGACTATAGGTGTACACCACACATCTGGATATTTTTT 2740
QY 1698 GTATTTTGTAGACACAGGGTTTCGCCATGTGCCAGGCTGGTCTTGAATTCCCTGAGC 1757
    |||||
Db 2739 GTATTTTGTAGAGACAGGGTTTCACCATGTTGCCAGGCTGGTCTTGAATGCCCTAGGC 2680
    |||||
QY 1758 TCAAGCAACCTGCCGCCTCGGCCTCCCAAGTACTGGGATTACAGCGAGAGGCCACCAT 1817
    |||||
Db 2679 TCATGCCATCTGCCTCGCCTCGGCCTACCAAGTTCTGGGATAACAAGCATGAGCCACTGT 2620
    |||||
QY 1818 GCCCAGGC 1825
    |||||
Db 2619 GCCCAGCC 2612
    |||||

```

Search completed: November 17, 2002, 16:40:59
 Job time : 268 secs